ASSOCIATION OF SOCIO-DEMOGRAPHIC CHARACTERISTICS WITH AUTISM SPECTRUM DISORDER: MEASUREMENT AND IDENTIFICATION TIMING IN THE U.S

by
Lydia A. Stewart

A dissertation submitted to Johns Hopkins University
in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
January 2019

ABSTRACT

Background: Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as a restrictive and repetitive pattern of behaviors or interests (APA, 2013). Early intervention is critical to maximizing the health and functional potential of children affected by ASD. Yet, many children are not diagnosed until school-age or later, minimizing their opportunity for early intervention. Limited research is available on how socio-demographic factors are associated with ASD measurement and identification timing in the U.S

Methods: This dissertation seeks to advance the understanding of how socio-demographic factors are associated with ASD measurement and the timing of identification in the U.S. Measurement variation across socio-demographic characteristics was examined using data from the Study to Explore Early Development (SEED). The performance and construct validity of a no-cost ASD measurement tool, the Ohio State Autism Rating Scale (OARS 12-item), was compared to current gold-standard ASD assessments in SEED. The association of socio-demographic factors with ASD identification timing was examined using data from the Autism and Developmental Disabilities Monitoring (ADDM) network.

Results: Analysis of OARS performance in SEED showed sensitivity and specificity above the recommended threshold for diagnostic accuracy across all socio-demographic subgroups. Lower OARS 12-item sensitivity and specificity among children scoring "below average" across Mullen Scales of Early Learning (MSEL) developmental domains suggest that alternative thresholds or subscales should be explored for children at lower developmental levels. The OARS 12-item demonstrated acceptable levels of

construct validity across the multiple methods used to examine it in this analysis. The strict measurement invariance found for all subgroups except poverty suggests that the OARS items are measuring underlying ASD constructs in similar ways across different groups of children.

In the ADDM network, median age at identification decreased from 6.3 to 5.3 years from SY 2006-SY 2012. The nested multivariate survival models showed that identification timing was significantly associated with racial/ethnic group, maternal education, child IQ, and study year. Non-substantial variation was found across poverty and study-site level variables.

Conclusions: ASD identification is associated with many of the same socio-demographic variables seen to influence other health disparities in the U.S. Further examination of how socio-demographic factors are related to disparities in ASD measurement and identification timing will help to inform and improve the current identification infrastructure in the U.S.

Academic and Dissertation Advisors:
M. Daniele Fallin, PhD, Professor and Chair, Department of Mental Health (Primary)
Li-Ching Lee, PhD, Associate Scientist, Department of Epidemiology

Dissertation Committee Members:
M. Daniele Fallin, PhD, Professor and Chair, Department of Mental Health
Janice Bowie, PhD, Professor and Chair, Department of Health Behavior and Society
Luke Kalb, PhD, Assistant Professor, Department of Mental Health
Alden Gross, PhD, Assistant Professor, Department of Epidemiology

Dissertation Committee Alternates:
Judith Bass, PhD, Associate Professor, Department of Mental Health
Anne Riley, PhD, Professor, Department of Population, Family and Reproductive Health

DEDICATION

This work is dedicated in honor of

D'Kai J. Vanlandingham

Wendy Klag Center for Autism and Developmental Disabilities Intern

THREAD Student

May your memory call us to recommit ourselves to the necessary and peaceful work of public health in our Baltimore community

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the many sources of support I have had during my time in the doctoral program. I am grateful for the mentorship support of faculty and staff at the Wendy Klag Center for Autism and Developmental Disabilities. Working with affiliated faculty as well as with the SEED and EARLI staff was an enriching experience and greatly helped to shape my research interests. The following WKC faculty and staff had a significant influence on this dissertation: Dr. Daniele Fallin, Dr. Li-Ching Lee, Dr. Heather Volk, Jamie Dahm, Michelle Landrum, Shenika Robinson and Nicole Williams. I would also like to acknowledge SEED and ADDM network collaborators on these analyses: Eric Rubenstein, Lisa Wiggins, Matthew Maenner, Russel Kirby, Maureen Durkin, Cordelia Rosenberg, Steven Rosenberg, Jennifer Hall-Lande, Bryn Harris, Gaye Windham, Eric Moody and Deborah Christensen.

I am also thankful for the support of several faculty members from across the School of Public Health including Dr. Judy Bass, Dr. Janice Bowie, and Dr. Rashelle Musci. I would like to acknowledge all members of the FBLV group that met weekly to collaborate and support each other as we engaged in different avenues of research related to ASD. I would like to further acknowledge the WKC funding and year of placement on the Children's Mental Health Services Training Grant which contributed to my success in the doctoral program.

I would also like to acknowledge and thank Dr. Barbara Howard and Dr. Raymond Sturner who, through part-time employment at Total Child Health, provided me with opportunities to engage in new areas of ASD-related research, mentorship, and financial stability throughout my time in the doctoral program. I would like to thank family and friends who provided much emotional support during this process. Finally, I would like to thank Shudong Wang and family, whose unwavering enthusiasm and support for "Miss Lydia" set me on this research path almost 7 years ago so that I might help make the world a more healthy, safe, and fair place for children with ASD to learn and grow.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Statement of Problem

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as a restrictive and repetitive pattern of behaviors or interests (APA, 2013). First recognized as a distinct psychiatric disorder in 1943, salient features of ASD documented by both Dr. Leo Kanner and Dr. Hans Asperger are now part of what is defined as ASD in the Diagnostic and Statistical Manual of Mental Disorders (DSM) version 5. The current estimated prevalence of ASD in the United States is 1 in 59 children, with a prevalence 4.5 times higher in males than females (Baio et al. 2018). The measured prevalence of ASD has increased steadily over the past several decades, but obtaining true global estimates of ASD prevalence is hindered by a lack of data from low and middle-income nations (Elsabagh et al. 2012). Within the U.S., disparities exist in the estimated prevalence of ASD among racial/ethnic groups and are linked to potential delays in identification or misclassification among African-American and Hispanic populations (Baio et al. 2018; Shattck et al. 2009; Mandell et al. 2007; Bergeer et al. 2009). Limited research is available on how socio-demographic factors are associated with ASD measurement and identification timing in the U.S. Bridging this gap in knowledge can help to inform solutions that increase access to screening and care and reduce disparities in age at identification and thus in access to early intervention.

The current pathway for the identification of ASD in the U.S. is a multi-step process involving community-based screening in pediatric primary care, specialist referral, and

multidisciplinary approaches for diagnostic evaluation. The American Board of Pediatrics recommends that all children be screened with ASD-specific tools at 18 and 24 months in conjunction with ongoing developmental screening and surveillance (Zwaigenbaum et al. 2015; Hagan et al. 2017). If the threshold for concern is met on the selected screening measure, such as the Modified Checklist for Autism in Toddlers (M-CHAT-R), community physicians refer families to seek formal evaluation from local developmental specialists. Commonly utilized developmental specialists include developmental pediatricians, psychologists, and speech language pathologists at hospital and community clinic settings. Programs funded through state Departments of Education, such as Infants and Toddlers for children younger than age 3, also provide limited initial evaluation in addition to early intervention services. Early interventions such as the Denver Early Start Model method of Applied Behavior Analysis (ABA) and Developmental, Individual Differences, and Relationship-based (DIR) approaches can reduce ASD-related social and communication impairments while helping children manage repetitive and restrictive behavior (Warren et al. 2011). When accessed within the critical developmental window between ages two and five, early intervention can help reduce impairment related to ASD and promote the social development necessary for later school success (Warren et al. 2011). Thus, early identification is critical to maximize potential for children affected by ASD.

Delays and disparities in the timely and appropriate identification of ASD can be attributed in part to *measurement* and *structural* barriers along the identification pathway. Barriers to ASD identification associated with measurement include variation in measurement tool performance across demographic and developmental subgroups, high

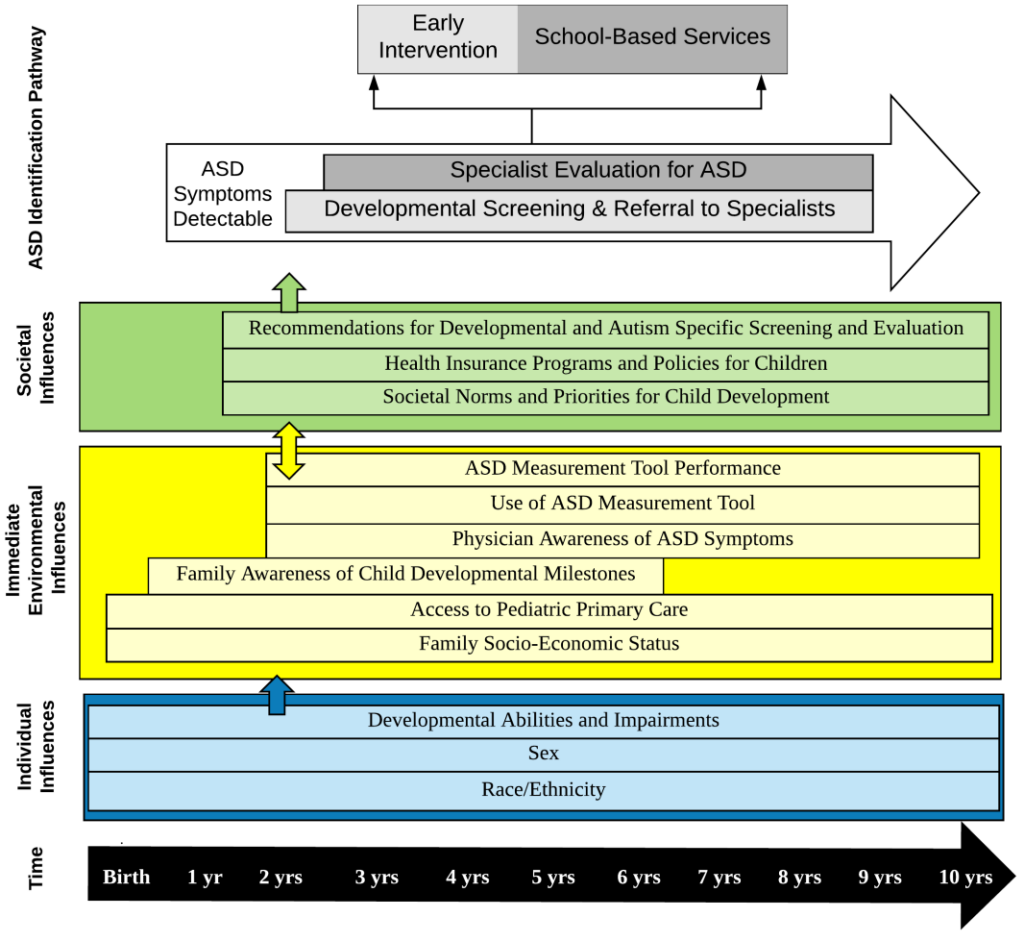
cost of per-use and other licensed tools, and administration methods that do not always fit the contexts or populations in which they are used (Durkin et al. 2015; Rosenberg et al. 2018; Khowaja, Hazzard & Robins, 2015; Stewart & Lee, 2017). Particularly in low resource communities, these factors may prevent or delay the standardized and multidisciplinary evaluation that is needed to access early intervention services.

The process from initial screening to receipt of diagnosis has been described by some caregivers as a "diagnostic odyssey" and can take anywhere from an estimated 2 months to over a year (Lappé, et al. 2018; Gordon-Lipkin et al, 2017; Bisgaier et al, 2011). Delays in identification have been attributed to structural barriers such as initial failure to screen for ASD in primary care, fragmented referral and care coordination systems, increasing numbers of children requiring evaluation, and a limited supply of clinical expertise and resources (Gordon-Lipkin et al. 2017; Bisgaier et al. 2011; Fenikile et al. 2015). These structural barriers to identification may disproportionally affect families of lower socio-economic status and members of racial/ethnic groups that already experience greater disparities in healthcare quality and access (Flores & The Committee on Pediatric Research, 2010; Gourdine et al. 2011; Becerra et al. 2014).

To guide this dissertation, a new conceptual model for examining influences on ASD identification was developed based on the Bronfrenbrenner ecological systems theory (Figure 1). A classic Bronfrenbrenner model nests individual characteristics within immediate environment influences such as family and community characteristics, indirect environmental influences such as government policies and cultural norms and a chronosystem to reflect change over time (Bronfrenbrenner & Morris 1998). Figure 1 disregards the Bronfrenbrenner nested spheres visual approach in favor of depicting

interacting levels of influence along a linear chronosystem. Figure 1 utilizes the time scale common to life-course theory models to better highlight critical windows for action along the ASD identification pathway. Individual level factors, such as developmental impairment and sex, influence and interact with factors in the immediate environment such as family awareness of developmental milestones and ASD measurement tool performance. The immediate environment in turn has a bi-directional relationship with societal influences, such as recommendations and policies for developmental screening in pediatric primary care and societal norms about typical child development at each age. In the model, all three of these levels interact at different points within the ASD identification pathway to influence timing of identification and access to intervention services.

Figure 1. Conceptual Model for the Social-Ecological Influences on Autism Spectrum Disorder Identification and Intervention in Early Childhood



This dissertation seeks to advance the understanding of how socio-demographic factors are associated with ASD measurement and the timing of identification in the U.S. Measurement variation across socio-demographic characteristics will be examined using data from the Study to Explore Early Development (SEED). SEED is a multi-site case-control study conducted at research institutions across the U.S since 2007 to gather information about ASD behavioral phenotype and associated medical, developmental and behavioral conditions as well as to examine possible environmental and genetic risk

factors for autism among children age two through five years (Schendel et al. 2012). In addition to the gold-standard, expensive and licensed diagnostic tools for ASD, the SEED protocol used the Ohio Autism Rating Scale (OARS 12-item). The OARS 12-item is a no-cost interactive measurement tool for ASD that uses brief clinical observation and interview with caregivers (OSU RPP, 2005). The OARS 12-item remains understudied as a measurement tool for assessing ASD among high-risk children. Yet, due to its open-source availability, it could be adapted for use as an additional tool in low resource settings. This dissertation examines OARS performance in SEED as a step towards its potential use as an alternative, no-cost, tool in low-resourced settings. The association of socio-demographic factors with ASD identification timing will be examined using data from the Autism and Developmental Disabilities Monitoring (ADDM) network, a collaborative, active surveillance system that provides estimates of the prevalence of ASD among children age 8 years (Baio et al. 2018). Building on prior work in the ADDM network, this dissertation examines the ages at ASD identification for children from the 2006, 2008, 2010 and 2012 ADDM study years (SY) and utilizes available census-linkage in the data to explore the association of identification timing with community level poverty alongside other socio-demographic factors.

1.2 Specific Aims

Aim 1: Compare the performance of the Ohio State University Autism Rating Scale (OARS) to gold-standard measures for determination of Autism Spectrum Disorder in children from the SEED I & II study.

This aim was foundational for the ASD measurement component of this project. The primary objective of this aim was to determine a threshold for OARS score that maximizes sensitivity and specificity across socio-demographic strata using ROC analyses. The performance of the OARS was also examined across developmental levels to investigate differences in measurement. Evaluation of OARS performance compared with gold-standard tools will allow for a proof of concept for OARS use as an ASD measurement tool for ages two-five years.

Hypothesis 1: The OARS will have sensitivity and specificity above 0.70 across socio-demographic strata.

Aim 2: Examine the construct validity of the Ohio State University Autism Rating Scale (OARS) for determination of Autism Spectrum Disorder in children from the SEED I & II study.

This aim examined the ability of OARS items to measure known ASD constructs. Multi-trait, multi-method matrices were constructed to determine the degree to which OARS items are associated with items of the same ASD constructs when measured by gold-standard assessments. Exploratory and Confirmatory Factor analyses were conducted to explore the underlying factor structure of the OARS and assess measurement invariance of items across socio-demographic strata. Evaluation of OARS construct validity through the multiple methods in this aim will determine whether OARS measures the known underlying ASD constructs in similar ways compared to other gold-standard assessments and across socio-demographic subgroups.

Hypothesis 2a: OARS items will cluster on a 2-factor structure.

Hypothesis 2b: OARS items will have strict measurement invariance across socio-demographic subgroups.

Aim 3a: Examine the variation in median age at identification across study years from 2006-2012 in the U.S. through the ADDM study.

Unadjusted and adjusted median age at identification was examined across birth cohorts using Kaplan Meier curves and multivariate survival analysis. Assessing variation in median age at identification across cohort years will contribute to the understanding of the ASD identification pathway in the U.S. When placed in the context of changes to clinical practice guidelines for community screening of ASD and public health initiatives to increase early identification and awareness, assessing changes in age at identification over time can help inform on the effectiveness of these initiatives.

Hypothesis 3a: Median age of identification will have significantly decreased from 2006-2012.

Aim 3b: To examine the variation in age at ASD identification by socio-demographic factors among children with ASD in the ADDM network.

Nested, multivariate survival analysis was used to examine the association of child, family, and surveillance network level factors on identification timing, clustered within geographically defined community level poverty and study sites. By examining how age at ASD identification has changed over time and what factors are related to

timing delays in the U.S, we can better inform and strengthen the current identification infrastructure.

Hypothesis 3: Controlling for child, family, and surveillance level factors, median age at identification will vary between high and low poverty communities.

1.3 Dissertation Outline

This dissertation is organized into six chapters. This introductory chapter provides an overview of the scope of the public health problem and brief rationale for studying the association of socio-demographic factors with ASD measurement and identification timing. Chapter 1 also serves an outline of the conceptual framework and specific aims of this dissertation.

Chapter 2 provides a review of relevant literature on current methods for identifying ASD in young children as well as discussions of disparities in ASD measurement and identification timing. It also provides a brief discussion of ASD measurement and identification timing concerns specific to low-resource settings within the U.S and globally and discusses current gaps in both knowledge and appropriate resources.

Chapters 3-5 report results for each specific aim through analyses conducted using data from the SEED and ADDM networks. They are formatted as stand-alone manuscripts for scientific publication. Chapter 3 reports on the performance of the OARS 12-item measurement tool compared to gold-standard ASD assessment methods across socio-demographic and developmental subgroups. Chapter 4 expands on the understanding of how the OARS measures ASD across socio-demographic and developmental subgroups by assessing the construct validity of the tool through the use

of factor analysis and measurement equivalence methods. Chapter 5 reports on changes in age at ASD identification from 2006-2012 and explores variation in identification timing by socio-demographic, developmental, and surveillance site level factors in the ADDM network.

Finally, Chapter 6 summarizes the findings from all results in relation to this dissertation's specific aims and hypotheses. This chapter also examines the public health implications of the results, future directions in the study of how socio-demographic factors are associated with ASD identification and discusses the relevance of these findings for improving ASD measurement and identification processes in low-resources settings.

1.4 References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (5th ed.). Arlington, VA: American Psychiatric Association.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C.), 67*(6), 1-23.

Becerra, T. A., Heck, J. E., & Olsen, J. (2014). Autism Spectrum Disorders and Race, Ethnicity, and Nativity: A Population-Based Study. *Pediatrics*, *134*(1), e63-e71.

Begeer, S., El Bouk, S., Boussaid, W., Terwogt, M. M., & Koot, H. M. (2009). Underdiagnosis and referral bias of autism in ethnic minorities. *Journal of Autism and Developmental Disorders*, *39*(1), 142.

Bisgaier, J., Levinson, D., Cutts, D. B., & Rhodes, K. V. (2011). Access to autism evaluation appointments with developmental-behavioral and neurodevelopmental subspecialists. *Archives of Pediatrics & Adolescent Medicine, 165*(7), 673-674.

Bronfenbrenner, U., & Morris, P. (1998). The ecology of developmental processes. In *Theoretical Models of Human Development* (pp. 993-1028). Wiley.

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... & Tager-Flusberg, H. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473-476.

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., ... & Yasamy, M. T. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, *5*(3), 160-179.

Fenikilé, T. S., Ellerbeck, K., Filippi, M. K., & Daley, C. M. (2015). Barriers to autism screening in family medicine practice: a qualitative study. *Primary Health Care Research & Development*, *16*(4), 356-366.

Flores, G., & Committee On Pediatric Research (2010). Technical report--racial and ethnic disparities in the health and health care of children. *Pediatrics*, *125*(4), e979.

Gourdine, R. M., Baffour, T. D., & Teasley, M. (2011). Autism and the African American community. *Social Work in Public Health*, *26*(4), 454-470.

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics, 63*(5), 851-859.

Hagan, J. F., Shaw, J. S., & Duncan, P. M. (2007). *Bright futures: Guidelines for Health Supervision of Infants, Children, and Adolescents* (4th ed.). Elk Grove Village, IL.: American Academy of Pediatrics.

Khowaja, M. K., Hazzard, A. P., & Robins, D. L. (2015). Socio-demographic barriers to early detection of autism: screening and evaluation using the M-CHAT, M-CHAT-R, and follow-up. *Journal of Autism and Developmental Disorders*, *45*(6), 1797-1808.

Lappé, M., Lau, L., Dudovitz, R. N., Nelson, B. B., Karp, E. A., & Kuo, A. A. (2018). The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics*, *141*(Supplement 4), S272-S279.

Mandell, D. S., Ittenbach, R. F., Levy, S. E., & Pinto-Martin, J. A. (2007). Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1795-1802.

OSU Research Unit on Pediatric Psychopharmacology (OSU RUPP). (2005). *OSU Autism Rating Scale-DSM 4 (OARS-4)*. Columbus, Ohio: OSU Research Unit on Pediatric Psychopharmacology.

Schendel, D. E., DiGuiseppi, C., Croen, L. A., Fallin, M. D., Reed, P. L., Schieve, L. A., et al. (2012). The study to explore early development (SEED): A multisite epidemiologic study of autism by the centers for autism and developmental disabilities research and epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders, 42*(10), 2121-2140.

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., & Baio, J. (2009). Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 474-483.

Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low-and middle income countries: A systematic review. *Autism*, *21*(5), 527-539.

Rosenberg, S. A., Moody, E. J., Lee, L. C., DiGuiseppi, C., Windham, G. C., Wiggins, L. D., ... & Young, L. (2018). Influence of Family Demographic Factors on Social Communication Questionnaire Scores. *Autism Research: official journal of the International Society for Autism Research*, *11*(5), 695.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-11.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., et al. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136 Suppl 1*, S41-59.

CHAPTER 2: REVIEW OF THE LITERATURE

2.1 ASD Identification Pathways in the U.S

As briefly discussed in Chapter 1, the current pathway for identification of ASD in the U.S. is a multi-step process involving community-based screening in pediatric primary care, specialist referral, and multidisciplinary approaches for diagnostic evaluation. While some parents may raise concerns about atypical development and delayed milestones before the first year of life, the average age of first parental concern for ASD ranges from ages 12-18 months in the scientific literature (Rosenberg et al. 2011; Young, Brewer & Pattinson, 2003; Hess & Landa, 2012; Wetherby et al. 2014). This average age of parental concern coincides with the American Academy of Pediatrics' 2007 recommendation for the use of ASD-specific tools at 18 and 24-month well-child visits in conjunction with ongoing developmental screening and surveillance (Johnson, Meyers, and AAP Council on Children with Disabilities, 2007; Zwaigenbaum et al. 2015; Hagan et al. 2017). Prior to 2007, the universal use of ASD-specific tools such as the Modified Checklist for Autism in Toddlers (M-CHAT-R) was not an AAP recommended part of routine care, rather for use only if specific developmental concerns were raised. A 2016 analysis of evidence on universal ASD screening in the U.S. by the U.S. Preventive Services Task Force (USPSTF) found insufficient evidence to make recommendations for or against the routine use of ASD screening for children below age 3, stating "this is not a recommendation against screening, it is a call for more research" (Siu & USPSTF, 2016). Primary concerns of the USPSTF were the lack of community-based studies examining screening performance in pediatric care and the efficacy of

subsequent referral of these screen positive children to evaluation and intervention. Without more evidence on the impact of referral and evaluation on children identified as part of universal screening, the task force was unable to adequately assess the balance of risk and benefit to early screening. As more research is conducted on the use of ASD-specific screening tools in routine pediatric care, a greater understanding of how these early-age measurement tools perform across socio-demographic and developmental subgroups is needed.

Once a child meets criteria for ASD risk on a screening tool in pediatric primary care and/or is perceived by a clinician as having sufficient developmental concerns to warrant further evaluation, referral is made to one or more developmental specialists (Zwaigenbaum et al. 2015; Hagan et al. 2017). Commonly used developmental specialists include developmental pediatricians, psychologists, pediatric neurologists, and speech language pathologists at hospital and community clinic settings. Many are now "autism centers" within university-affiliated hospitals where evaluation is performed by a multidisciplinary team of specialists (Zwaigenbaum et al. 2015; Hagan et al. 2017; Huerta & Lord, 2012). Programs funded through state Departments of Education, such as Infants and Toddlers for children younger than age 3, also provide limited initial evaluation in addition to early intervention services. Once general developmental and autism-specific evaluation has identified areas of developmental concern or impairment sufficient for a diagnosis, children are considered eligible for intervention services as mandated by the Individuals with Disabilities Education Act (IDEA of 2004). State-funded early intervention programs such as Infants and Toddlers target children younger than age 3 and have varying eligibility criteria across states with regard to percentage and

types of developmental delay diagnosed (ECTA "States' and Territories' Definition of/Criteria for IDEA Part C Eligibility", 2015). After age 3, children are eligible for appropriate services overseen and administered by local Departments of Education guided by an individualized education plan (IEP).

2.2 Instruments for the Measurement of ASD

The previously outlined ASD identification pathway uses multiple types of screening and evaluation measures. A *screening* measure is a clinical test or procedure performed on members of a defined asymptomatic population or population subgroup to assess the likelihood of their members having a particular disorder (Maxim, Niebo & Utell, 2014). Screening measures are not intended to diagnosis a disorder, rather to be used in sequential testing to efficiently identify individuals in need of further evaluation.

In pediatric primary care, brief screening measures such as the M-CHAT-R/F are used for identifying children with potential ASD from the general population (Robins et al. 2014). More complex screening measures, such as the Social Responsiveness Scale (SRS) and Social Communication Questionnaire (SCQ), can require more time to complete and expertise to score and are often used in primary care and specialist settings for children already identified as high-risk for ASD (Bridgemohan et al. 2018; Stewart & Lee, 2017). These screening instruments have items designed to identify ASD-associated behaviors based on child age and are completed as caregiver self-complete in a written format, a common method of administration for screening instruments. Other screening tools such as the Screening Tool for Autism in Toddlers and Young Children (STAT) and the Parent Observation of Social Interaction (POSI) employ clinician or parent

observation and interaction with children and to assess ASD risk (Stone et al, 2004; Stone et al, 2013; Smith, Sheldrick & Perrin, 2013). On screening instruments, ASD behaviors are often either endorsed as a binary "yes/no" response or measured in terms of frequency or impairment of the behavior on a likert scale from "rarely" or "not a problem" to "often" or "a significant problem". The SCQ and SRS are licensed through Western Psychological Services (WPS) and are priced ranging from $2-5 per use in addition to initial training materials (WPS "Pricing Information"). The STAT is licensed through Vanderbilt University and is currently listed as $500 USD per administration kit plus $200 USD per STAT training certificate (Vanderbilt University E-Innovations "STAT").

After screening has identified children at higher risk for ASD, *assessment* measures are used to evaluate specific symptom profiles, severity, and provide a diagnosis (Maxim, Niebo & Utell, 2014). ASD evaluation measures used in diagnosis rely on direct observation and clinical interaction with children or caregivers to gain a broad understanding of the child's development and behavior profile. The Autism Diagnostic Observation Schedule (ADOS) is often considered the "gold-standard" measure of ASD in children (Lord et al. 2000). The ADOS is a standardized diagnostic measure that involves direct interaction and scoring of child's behavior, taking into account developmental level and age of the child (Lord 2012). Lasting 45-60 minutes, the assessment includes standardized evaluator-administered activities that engage the child in tasks that prompt social interactions, communication, and repetitive behaviors for examination. The ADOS has four modules based on chronological age as well as developmental and language levels. Modules 1 and 2 pertain to children with 48 months language development, while modules 3 and 4 are appropriate for older children and

adolescents. ADOS items cover a range of social communication items such as “response to joint attention” and “facial expression” as well as items measuring restricted and repetitive patterns of interest or behavior such as “unusual sensory interests” (Gotham et al. 2007). Licensed through WPS, an ADOS-2 administration kit currently costs $2,095 USD and purchasing is restricted to individuals who have either advanced educational qualifications such and appropriate professional licensure and certification (WPS “Pricing Information”). In early validation testing by the ADOS authors, modules 1, 2, and 3 had sensitivities and specificities of (module 1) 97% & 95%, (module 2) 95% & 87%, and (module 3) 90% & 94% respectively (Lord et al. 2000). Previous analysis in the SEED 1 sample population found a sensitivity, specificity and positive predictive value of the ADOS to be 91%, 82%, and 80% respectively (Wiggins et al. 2015). Some studies have found variation in ADOS performance across sex, race, and developmental level (Bishop et al. 2017; Harrison et al. 2017; de Bildt et al. 2004; Gotham et al. 2007)

The Autism Diagnostic Observation Schedule (ADI-R) is an additional “gold standard” evaluation measure for ASD (Lord, Rutter & Couter et al. 1994) The ADI-R is a semi-structured parent/caregiver interview that includes 93 questions about ASD characteristics across the three DSM-IV-TR domains of language/communication, reciprocal social interaction and restricted, repetitive and stereotyped behaviors and interests. All ADI-R items are coded for past and current behavior. Scoring is coded using different algorithms based on age and can be administered to caregivers of children as young as 12 months (de Bildt et al. 2015). The ADI-R is licensed through WPS and an initial starter kit ,including interview booklets and ten scoring sheets, costs $275 USD with additional per-use priced scoring sheets available for purchase separately (WPS

"Pricing Information") The ADI-R has a reported sensitivity and specificity of over 90% (Lord et al. 1993). Prior work in SEED I found an ADI-R sensitivity of 77% and specificity of 73% (Wiggins et al. 2015). ADI-R performance has been shown to vary by sex and race/ethnicity (Tillman et al. 2018; Vanegas et al. 2016).

The Mullen Scales of Early Learning (MSEL) is a developmental assessment often used to evaluate developmental delays in concert with ASD specific assessment measures. The MSEL is comprised of five subscales: gross motor, fine motor, visual reception, expressive language and receptive language (Mullen, 1995). Completion of a variety of progressively more difficult tasks in the 5 domains produces an early learning composite score as well as 5 domain sub-scores. Corresponding t-scores are selected based on sub scores and child age at evaluation, accounting for prematurity. An early learning composite score is calculated based on each domain t-score. Ratings on the MSEL have been found to have internal consistency ranging from 0.75-0.83 as well as strong correlation with other measures of the same developmental domains such as the Preschool Language Assessment and the Peabody Fine Motor Scale (Shank, 2011). The MSEL, like other tools mentioned previously, is a licensed tool through WPS. An MSEL starter kit including test materials, scoring manual, 25 scoring sheets, and item administration book currently costs $956.80 USD with additional score sheets available at per-use pricing (WPS "Pricing Information").

The Ohio State Autism Rating Scale (OARS 12-item) is a semi-structured interview administered scale for clinicians and trained researchers with three DSM-IV based subscales for social impairment, communication impairment and restricted, repetitive and stereotyped behavior (see appendix 1). Importantly, the OARS is

completely free for use. The OARS has 12 items rated on an ordinal scale of "0: Never or Rarely; not a problem" to "3: Very Often; a severe problem". The OARS obtains measures of ASD severity through a calculation of two scores: "Impairment Mean" and a "Total Symptom Count". Symptom count for each domain and for total symptom count is calculated by adding the number of symptoms that received a score of greater than "0: Rarely or Never". The impairment mean for each domain and total impairment mean are calculated by adding the number of points 0-3 gained for severity of each item then dividing by the number of points possible. The denominators for verbal children are 12 for social interaction, 12 for communication, 12 for restricted patterns. The denominators for nonverbal children are 12 for social interaction, 9 for communication (excluding item B2), and 12 for restricted patterns (OSU RPP, 2005). Two related autism scales published online in the same public-use document as the OARS include the "OSU Global Severity Scale for Autism" and "OSU Global Improvement scale" (OSU RPP; 2005). The related "OSU Global Severity Scale for Autism", often referred to in literature as the OARS CGI, has been shown to have acceptable validity and reliability in prior studies (Olson & Bolte, 2013; Olson et al, 2017). At this time, no formal validation of the 12-item OARS has been conducted to compare OARS performance and measurement properties to gold-standard ASD assessment tools (OSU RPP, 2005; Stewart & OSU, Personal communication, June 2018).

2.3 Socio-demographic Variation in ASD Measurement and Identification

A central theme of this dissertation is the examination of how ASD measurement and identification timing vary by socio-demographic factors such as sex, race/ethnicity,

and socio-economic status. In this section, a brief overview of the literature is presented on several socio-demographic factors that will be examined in chapters 3-5.

Autism is four times more prevalent in males than females (Baio et al. 2018). Several studies have found a slight but statistically significant earlier age of concern (AOC) among parents of female children compared to males, possibly due to earlier initiation of social communication in female children (Rosenberg et al. 2011; De Giacomo, 1998). Despite earlier AOC for female children, previous work in the ADDM network found a median age of identification of 6.1 years for females compared with 5.6 for males (Shattuck et al. 2009). One potential explanation of this delay may be the occurrence of symptom "masking" whereby female children are able to better integrate verbal and non-verbal behaviors, maintain reciprocal conversations, and exhibit a different pattern of restricted and repetitive interests than their male peers (Hiller et al. 2014; Mandy et al. 2012). Female children with ASD have also been shown through studies of teacher report to exhibit fewer externalizing behaviors than male children of the same age (Hiller et al. 2014; Mandy et al. 2012).

A female presentation of ASD with fewer externalizing behaviors and "masking" of deficits in social communication differently than males of the same age presents a measurement challenge for tools developed and validated in predominantly male sample populations. While some instruments such as the Social Responsiveness Scale have not been reported to have statistically significant variation in performance by sex, studies on the ADOS and ADI-R have shown significant differences in the performance of measured restrictive and repetitive behavior domains between males and females, resulting in lower sensitivity (Werling & Gerschwind, 2013; Tillman et al. 2018; Wang et

al. 2017). Potential differences in the developmental trajectory of ASD in females as compared to males is an emerging area of focus in the ASD research community. Future research should consider differences in ASD presentation by sex during measurement instrument development and validation.

In the U.S, there is a higher measured prevalence of ASD among non-Hispanic white children (17.2 per 1,000) compared with their non-Hispanic African-American (16.0 per 1,000), Asian/Pacific Islander (13.5 per 1,000) and Hispanic peers (14.0 per 1,000) (Baio et al. 2018). Studies have found contradictory results with regards to the statistical significance of race/ethnicity associated with ASD identification timing. Previous work in the ADDM network using the 2002 surveillance year cohort found that after adjusting for other covariates such as child IQ, maternal education, and study site, there was not a statistically significant variation in ASD identification timing by race/ethnicity (Shattuck et al. 2009). The latest report on ASD estimated prevalence in the ADDM network for the 2016 cohort found no significant difference in unadjusted median age at ASD identification by race/ethnicity (Baio et al. 2018). However, the examination of age of diagnosis of ASD in Philadelphia Medicaid claims data, the National Survey for Children with Special Healthcare Needs, and parent-report in the Interactive Autism Network all have found a statistically significant association between racial/ethnic groups other than white, non-Hispanic and age at identification (Mandell, 2002; Jo et al. 2015; Rosenberg et al. 2011).

With regard to race/ethnicity, literature suggests that differential measurement of ASD across subgroups is due to a combination of several factors including potential differences in symptom profiles and culture-based biases in measurement instruments

(Harrison et al. 2007). A community-based study from Philadelphia found that white, non-Hispanic children were more likely to exhibit inflexible adherence to routines and rituals and preoccupation with parts of a toy or object than their African-American peers (Sell et al. 2012). A population-based study from Los Angeles County, California found that African-American and both foreign born and U.S-born Hispanic children were more likely to exhibit an ASD presentation with severe emotional outbursts and impaired expressive language than the white, non-Hispanic children in the sample (Becerra et al. 2014). A separate study of children with an ASD diagnosis in Philadelphia Medicaid data found that African-American children were 2.6 times less likely to receive a diagnosis of ASD at first specialty visit and were significantly more likely to receive adjustment and conduct disorder diagnoses than white children (Mandell et al. 2007).

The evidence discussed above suggests possible variation in ASD presentation, but it also requires that we examine sources for potential biases in the measurement and clinical perception of ASD across racial/ethnic subgroups. ASD-specific instruments were designed to identify social communication and restricted and repetitive behavioral patterns that deviate from a normative set of developmental expectations for "typical" development at a given age. However, the development of these measurement tools for ASD in samples that predominantly include white, male children may lead to instruments that lack cross-cultural validity and fail to adequately detect the developmental social norms of other racial/ethnic communities (Harrison et al 2007). The classic example of cross-cultural variation in the types of social communication measured by ASD instruments considers the use of eye contact during social interactions. Multiple studies have found that perceptions of the importance of eye contact during communication

varies across cultures, with some cultural traditions such as those from Kenya and Japan discouraging direct eye contact between children and adult authority figures as a sign of respect (Harrison et al. 2007; McCarthy et al. 2006; Carter et al. 2005; Grinker, 2008). Cross-cultural variations have also been found in the use of nonverbal communication forms like facial expression, key elements of verbal communication such as conversational turn-taking, and play activities (Harrison et al. 2007; Elfenbein et al. 2007, Elfenbein, 2013; Marsh et al. 2003; Yuki et al. 2007; Carter et al. 2005). As the cultural adaptation of existing ASD tools and development of wholly new measures increases in countries that have historically not been a part of the creation of ASD measurement models, we will gain a greater understanding of cultural variability in ASD related behaviors.

In considering differences in ASD presentation, it is critical to also consider the environments, or contexts, in which children are living. Child development is shaped by a combination of individual factors, such as genetic predisposition, and environmental factors, such as increased exposure to harmful substances like lead and limited access to material and social resources (Letourneau et al. 2013; Bradley et al. 2002). Therefore, any examination of variation in ASD measurement or identification timing must also examine the socio-economic conditions known to shape child development and contribute to disparities in healthcare quality and access. In line with broader research on health disparities and diagnosis timing, family and community-level socio-economic status have emerged as a factor influencing ASD identification timing. Analyses using the National Survey on Children with Special Healthcare Needs found that children living in households with self-reported incomes <200 % of the federal poverty level (FPL) were

44% more likely to have a later diagnosis than their high income peers (Jo et al. 2015). The diagnosis delay was quantified by a later study looking specifically within the U.S state of Pennsylvania. In this sample, children living <100% of the FPL received an ASD diagnosis almost a full year later than their higher income neighbors (Mandell et al. 2005). Family and community SES is associated with access to medical care, with lower income populations experiencing greater geographic and financial barriers to access (Letourneau et al. 2013). The associations seen between poverty and ASD identification delay may represent this lack of access to early screening and diagnosis, but factors related to low-income such as lower maternal educational attainment, greater chronic disease burden, and limited health literacy may also play a role.

2.4 Research Gap in Low Resource Settings

A review of ASD under DSM-4 criteria reported a median global prevalence of 62 cases per 10,000 individuals, but lacked substantial representation of low and middle-income nations (Elsabagh et al. 2012). In addition to impeding the ability of researchers to make cross-country comparisons and learn about etiology in diverse populations, disjointed ASD identification infrastructure in low resource settings delays the provision of interventions that can reduce impairment. Conducting epidemiologic studies and surveillance of ASD requires the use of validated screening and diagnostic tools and efficient methods that accurately reflect the unique cultural and resource needs of the community.

A recent review of the use of ASD screening in low and middle-income nations found wide variation in study design, screening tools, and screening methodology

(Stewart & Lee, 2017). Few ASD screening tools designed to identify ASD risk in the general population have been validated for use outside of the high-income nations in which they were first developed.

A joint statement by leaders in the ASD research community from 2015 calls for the development of more open-source ASD assessment tools and technologies, greater cross-cultural validation of existing tools, and innovative approaches to training professionals and paraprofessionals in low-resource settings (Durkin et al. 2015). As the autism research community commits to addressing ASD burden in low-resource communities, a strong evidence-base on screening and evaluation tool performance in diverse populations will be needed. Specific considerations for the use of ASD measurement tools in low-resource communities and implications of this dissertation's findings for low-resource communities are discussed further in Chapter 6.

2.5 References

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371-399.

Bishop, S.L. Sweeney, M.P Huerta, M. Havdahl, A. Lord, C. (2017) Sex Differences in ADOS-2 Scores and Classifications among Verbally Fluent Children with Autism Spectrum Disorders. International Meeting for Autism Research. San Francisco, California. Retrieved from https://insar.confex.com/insar/2017/webprogram/Paper25811.html

Bridgemohan, C., Augustyn, M., & Torchia, M. M. (2018). Autism spectrum disorder: Screening tools. *UpToDate, Waltham, MA.* Accessed August, 2018.

Carter, J. A., Lees, J. A., Murira, G. M., Gona, J., Neville, B. G., & Newton, C. R. (2005). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language & Communication Disorders*, *40*(4), 385-401.

Constantino JN and Gruber CP (2005) *Social Responsiveness Scale: Manual*. Los Angeles, CA: Western Psychological Services.

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child & Adolescent Psychiatry*, *7*(3), 131-136.

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... & Tager-Flusberg, H. (2015). Autism screening and diagnosis in low resource settings: challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473-476.

Early Childhood Technical Assistance Center (2015) *States' and Territories' Definition of/Criteria for IDEA Part C Eligibility.* Retrieved from http://ectacenter.org/topics/earlyid/partcelig.asp

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, *5*(1), 90-96.

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion*, *7*(1), 131.

Grinker, R. R. (2008). *Unstrange minds:Remapping the world of autism*. Basic Books (AZ).

Harrison, A. J., Long, K. A., Tommet, D. C., & Jones, R. N. (2017). Examining the role of race, ethnicity, and gender on social and behavioral ratings within the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, *47*(9), 2770-2782.

Hess, C. R., & Landa, R. J. (2012). Predictive and concurrent validity of parent concern about young children at risk for autism. *Journal of Autism and Developmental Disorders*, *42*(4), 575-584.

Hiller, R.M., Young, R.L. & Weber, N. (2014) Sex Differences in Autism Spectrum Disorder based on DSM-5 Criteria: Evidence from Clinician and Teacher Reporting. *Journal of Abnormal Child Psychology 42,* 1381-1393.

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric Clinics of North America*, *59*(1), 103.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Johnson, C. P., Myers, S. M. & Council on Children with Disabilities (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1183-1215.

Letourneau, N. L., Duffett-Leger, L., Levac, L., Watson, B., & Young-Morris, C. (2013). Socioeconomic status and child development: A meta-analysis. *Journal of Emotional and Behavioral Disorders*, *21*(3), 211-224

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205-223.

Mandell, D. S., Listerud, J., Levy, S. E., & Pinto-Martin, J. A. (2002). Race differences in the age at diagnosis among Medicaid-eligible children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(12), 1447-1453.

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics*, *116*(6), 1480-1486.

Mandell, D. S., Ittenbach, R. F., Levy, S. E., & Pinto-Martin, J. A. (2007). Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1795-1802.

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: evidence from a large sample of children and adolescents. *Journal of Autism Developmental Disorders*, 42 (7), 1304-1313.

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal "accents" cultural differences in facial expressions of emotion. *Psychological Science*, *14*(4), 373-376.

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: a review with examples. *Inhalation toxicology*, *26*(13), 811-828.

McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2006). Cultural display rules drive eye gaze during thinking. *Journal of Cross-Cultural Psychology*, *37*(6), 717-722.

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, *133*(1), 37-45.

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research and Treatment*, *2011* Article ID 874619.

Shank L. (2011) Mullen Scales of Early Learning. In: Kreutzer J.S., DeLuca J,. Caplan B. (eds) *Encyclopedia of Clinical Neorpsychology*. Springer, New York, NY.

Siu, A. L., Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., Ebell, M., ... & Krist, A. H. (2016). Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *Journal of the American Medical Association*, *315*(7), 691-696.

Smith, N. J., Sheldrick, R. C., & Perrin, E. C. (2013). An abbreviated screening instrument for autism spectrum disorders. *Infant Mental Health Journal*, *34*(2), 149-155.

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of autism and developmental disorders*, *34*(6), 691-701.

Stone, W. L. (2013). Screening tool for autism in two-year-olds (STAT). *Encyclopedia of Autism Spectrum Disorders*, 2684-2688.

Tillmann, J., Ashwood, K., Absoud, M., Bölte, S., Bonnet-Brilhault, F., Buitelaar, J. K., ... & De Bildt, A. (2018). Evaluating sex and age differences in ADI-R and ADOS scores in a large European multi-site sample of individuals with autism Spectrum disorder. *Journal of Autism and Developmental Disorders*, 1-16.

Vanegas, S. B., Magaña, S., Morales, M., & McNamara, E. (2016). Clinical validity of the ADI-R in a US-based Latino population. *Journal of Autism and Developmental Disorders*, *46*(5), 1623-1635.

Vanderbilt University E-Innovations "Screening Tool for Autism in Toddlers and Young Children STAT" (2018) Retrieved from http://stat.vueinnovations.com

Wang, S., Deng, H., You, C., Chen, K., Li, J., Tang, C., ... & Zou, X. (2017). Sex differences in diagnosis and clinical phenotypes of Chinese children with autism spectrum disorder. *Neuroscience bulletin*, *33*(2), 153-160.

Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, *26*(2), 146.

Western Psychological Services (2017). “Pricing Information: 2018 Catalog”. Retrieved 11 November 2018 from https://www.wpspublish.com/store/2019-WPS

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autsim and Developmental Disorders*, *34*(5), 473-493.

Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., ... & Levy, S. E.(2015). Using standardized diagnostic instruments to classify children with autism in the study to explore early development. *Journal of Autism and Developmental disorders*, *45*(5), 1271.

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, *43*(2), 303-311

Young, R. L., Brewer, N., & Pattison, C. (2003). Parental identification of early behavioural abnormalities in children with autistic disorder. *Autism*, *7*(2), 125-143.

CHAPTER 3. PERFORMANCE OF THE OHIO STATE AUTISM RATING SCALE AS A POTENTIAL TOOL FOR IDENTIFYING AUTISM SPECTRUM DISORDER USING DATA FROM THE STUDY TO EXPLORE EARLY DEVELOPMENT

3.1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as a restrictive and repetitive pattern of behaviors or interests (APA, 2013). The current estimated prevalence of ASD in the United States is 1 in 59 children, with a significantly higher prevalence in males than females (Baio et al. 2018). The latest work from the Autism and Developmental Disabilities Monitoring Network reports a median age at identification of 4.3 years in the U.S, with variation in identification timing between sexes and across geographic region (Baio et al. 2018). Previous studies in clinical and community-based samples have also found variation in age at identification by racial/ethnic subgroup and household income (Shattuck et al. 2012; Mandel et al. 2002; Jo et al. 2015).

Early identification of ASD is key to providing early intervention services that may reduce ASD-associated impairment (Warren et al. 2011). Many early intervention frameworks have been designed to address core features of ASD during the critical developmental window between 2-5 years. Early interventions that feature Applied Behavior Analysis (ABA) and Developmental, Individual Differences, and Relationship-based (DIR) approaches can reduce ASD-related social and communication impairments while helping children better manage repetitive and restrictive behavior (Dawson et al. 2010; Warren et al. 2011). When implemented as part of comprehensive early intervention in ages 2-5, these therapies help prepare children with ASD to integrate more fully into the classroom and engage with peers.

In the US, the American Academy of Pediatrics recommends the use of ASD-specific screening instruments during well-child visits at ages 18 and 24 months in conjunction with ongoing developmental surveillance and screening (Zwaigenbaum et al. 2015; Hagan et al. 2017). Early identification of ASD through screening is critical for enabling timely physician referral to diagnostic specialists (Robins 2016; Koegel et al. 2005). In general pediatric primary care, brief, low-cost level-one screening measures such as the Modified Checklist for Autism in Toddlers-Revised (M-CHAT-R) are used for identifying a sample at high risk of ASD based on behaviors. Level-two screening measures, such as the Screening Tool for Autism in Toddlers and Young Children (STAT), can require more time and expertise to complete and are used in primary care and specialist settings for children already identified as high-risk for ASD (Stone et al, 2004; Stone et al, 2013). Once the threshold for concern is met on the selected screening measure, community physicians can refer families to seek formal evaluation from local developmental specialists.

Despite these recommendations, the average age of ASD diagnosis in the US is around 4 years, limiting opportunities for very early intervention. This delayed identification may be attributable to lack of universal screening, inadequate or inappropriate screening tools, limited affordability of tools, or the 'diagnostic odyssey' in which it takes years for a child to receive a formal diagnosis after concern has been raised (Baio et al, 2018; Jo et al. 2015; Gordon-Lipkin, Foster, and Peacock, 2016).

The lack of identification or delay in time from identification to diagnosis may especially affect families of lower socio-economic status. Depending on a family's geographic location, insurance coverage, ability to pay out-of-pocket, and language

spoken, the wait time for a recommended comprehensive ASD evaluation can take from 2 months to a year (Gordon-Lipkin et al, 2017; Bisgaier et al, 2011; Astin et al, 2015). In communities with already overburdened clinical infrastructures, the limited supply of developmental subspecialists with training on ASD evaluation, high cost of per-use licensed screening and assessment materials, and recent increase in the number of children needing early ASD evaluation and referral can lead to delays in providing appropriate ASD identification (Gordon-Lipkin et al. 2017; Warren et al. 2009; Schulz, 2016).

To address inadequate and delayed identification of ASD, more research is needed on new or under-researched measures and processes that complement existing efforts and are affordable for low-resource communities. One under-researched tool that could be useful in this context is the Ohio State Autism Rating Scale (OARS 12-item). Published for public use in 2005 and based on DSM-IV ASD domains, the OARS 12 item is a no-cost, interactive tool that uses clinician observation as well as semi-structured interview with the primary care giver (OSU RPP, 2005). The OARS authors report an average 30 minute administration time. This brief scale measures the three DSM-IV domain criteria through a series of 12 questions rated on a four level ordinal scale of "Never or Rarely; not a problem" to "Very Often; a severe problem". The OARS obtains measures of ASD severity through a calculation of two scores: "Total Impairment Mean" (TIM) and a "Total Symptom Count" (TSC). Given the short assessment time, and the affordability of OARS, it may be a useful alternative to higher cost tools in busy low-resourced communities. To our knowledge, no validation of the 12-item OARS compared to gold-standard assessments has been conducted (OSU RPP, 2005; Olson &

Bolte, 2013; Olson et al, 2017). The addition of the no-cost OARS as a level-two screener alongside standard care may help to offset the higher volume of need and reduce delays to obtaining sufficient evaluation to access developmental specialists in low resource settings.

Our objective was to assess the performance of the OARS 12 item in phases I and II of the Study to Explore Early Development (SEED) by comparing its ASD classification performance to gold-standard instruments and diagnosis in a clinic-based diagnostic setting among children at high-risk. Procedures for classification of ASD in the SEED population limits the extent to which formal validation of OARS as a screening tool can be accomplished because OARS was implemented as a part of the clinical evaluation rather than as a screening tool in SEED. Thus, clinicians had knowledge of child symptoms and behaviors when completing the OARS. However, the availability of both OARS and gold-standard ADOS and ADI-R data on the same children in SEED does provide a proof of concept opportunity to assess the plausibility of the OARS 12-item as a method of identifying ASD symptoms in a high-risk sample. We believe that examining the performance of OARS by socio-demographic differences in SEED may enable future avenues for adaptation and implementation of the OARS as an additional measurement tool in low resource areas.

3.2 Methods
Participants

SEED is a multi-site case-control study conducted in multiple phases at research institutions across the U.S since 2008 (Schendel et al. 2012; Centers for Disease Control and Prevention, 2018). SEED was established to gather information about ASD

behavioral phenotype and associated medical, developmental and behavioral conditions as well as to examine possible environmental and genetic risk factors for autism among children age 2-5 (Schendel et al. 2012; Wiggins et al, 2015). Participants were recruited from sites in California, Colorado, Georgia, Maryland, North Carolina, and Pennsylvania between the ages of 30-64 months. SEED phase 1 recruited children from 2007-2011 and phase II from 2012-2016.

Typically-developing children were recruited from the general population based on a random sampling of state vital records. Children with potential developmental delay (DD) or ASD were invited to participate after being identified from educational and healthcare providers that serve children with DD. Children were excluded from participation in the SEED study if they were born outside the designated study catchment areas, had mobility, vision or hearing impairments that prevented full engagement in the assessments, or did not meet spoken language requirements. Participants with a prior ASD diagnosis or who scored above an 11 (Wiggins et al 2015) on the Social Communication Questionnaire (Rutter, Bailey and Lord, 2003) in the first phone interview went on to receive an Autism evaluation in addition to a developmental assessment and collection of bio-samples. Participants not meeting these requirements received only developmental assessment and bio-sample collection. During the autism evaluation component, autism specific parent interviews and child observation measures were performed. Participants without complete data on the OARS tool or who did not receive a final case determination were excluded from this analysis.

Developmental Measures

Autism Diagnostic Observation Schedule-2 (ADOS)

The ADOS is a standardized diagnostic measure that involves direct interaction and scoring of child's behavior, taking into account developmental level and age of the child (Lord 2012; Lord et al. 2000). ADOS items cover a range of social communication items such as "response to joint attention" and "facial expression" as well as items measuring restricted and repetitive patterns of interest or behavior such as "unusual sensory interests" (Gotham et al. 2007). Prior analyses of the ADOS in the SEED 1, using final case status and the gold standard, found sensitivity and specificity of 0.91 and 0.62 respectively (Wiggins et al. 2015b).

Autism Diagnostic Interview-Revised

The ADI-R is a semi-structured parent/caregiver interview that includes 93 questions about ASD characteristics across the three DSM-IV-TR domains of language/communication, reciprocal social interaction and restricted, repetitive and stereotyped behaviors and interests (Lord et al. 1994). All ADI-R items are coded for past and current behavior. Scoring is coded using different algorithms based on age and can be administered to caregivers of children as young as 12 months (de Bildt et al. 2015). The ADI-R has a reported sensitivity and specificity of over 90% (Lord et al. 1993). Prior work in SEED I found an ADI-R sensitivity of 77% and specificity of 73% (Wiggins et al. 2015).

Mullen Scales of Early Learning (MSEL)

The MSEL is a developmental assessment comprised of five subscales: gross motor, fine motor, visual reception, expressive language and receptive language (Mullen, 1995). Completion of a variety of progressively more difficult tasks in the 5 domains produces an early learning composite score as well as 4 domain sub-scores. Corresponding t-scores are selected based on sub scores and child age at evaluation. An early learning composite score is calculated based on each domain t-score. In this analysis, a t-score below 85 on the early learning composite score was considered "below average".

Ohio State Autism Rating Scale (OARS 12 item)

The OARS 12-item is a semi-structured interview administered scale for clinicians and trained researchers with three DSM-IV based subscales for social impairment, communication impairment and restricted, repetitive and stereotyped behavior. The OARS has 12 questions rated on an ordinal scale of "0: Never or Rarely; not a problem" to "3: Very Often; a severe problem". The OARS obtains measures of ASD severity through a calculation of two scores: "Impairment Mean" and a "Total Symptom Count". Symptom count for each domain and for total symptom count is calculated by adding the number of symptoms that received a score of greater than "0: Rarely or Never". The impairment mean for each domain and total impairment mean are calculated by adding the number of points 0-3 gained for severity of each item then dividing by the number of points possible. The denominators for verbal children are 12 for social interaction, 12 for communication, 12 for restricted patterns. The denominators

for nonverbal children are 12 for social interaction, 9 for communication (excluding item B2), and 12 for restricted patterns (OSU RPP, 2005).

Final Case Status

Final case determination of ASD and DD status was made using a previously published algorithm (Wiggins et al 2015) based primarily on the ADI-R and ADOS. If the child met SEED ASD criteria but had a mental age less than 24 months, the OARS degree of certainty (DOC) rating (added for SEED) was used to determine case status (Wiggins et al. 2015). Specifically, if clinicians were certain the child had ASD (i.e. DOC rating of 4 or 5) the child was classified as an ASD case and if clinicians were uncertain the child had ASD (i.e., DOC rating of 1, 2, or 3) the child was classified as "indeterminate." Indeterminate cases were subsequently dropped from analyses because ASD could not be confirmed, or could not be distinguished from intellectual disability.

Maternal Interview

Mothers reported socio-demographic factors via SEED questionnaires. Of particular interest for this study were maternal and paternal race/ethnicity, maternal educational attainment, maternal employment patterns during pregnancy and estimated household number of residents and income during pregnancy and at time of study enrollment. Income level relative to the Federal Poverty Level (FPL) was calculated using guidelines and thresholds for 2012 and accounted for number of persons in household as well as income for the year prior to the child's evaluation (HHS, "Poverty Guidelines" 2012).

Analytic sample

For this analysis, we included all children enrolled in SEED I or II who received a full diagnostic evaluation for ASD, including the OARS. The analytic sample was restricted to only children who did not have discordance between ADOS and ADI-R, meaning that the OARS 12 item related OSU CGI scale was not used in determining final case status. Children were also excluded from the analytic sample if they did not have complete data on the OARS-IV measure or lacked a determinate final case.

Statistical analysis

Descriptive Statistics: To address the small sample size across socio-demographic subgroups, participants were analyzed as ASD and Non-ASD. The non-ASD combined the Population and DD subgroups who had the full developmental evaluation into a single group, with 83% of participants in this category originally belonging to the Developmental Delay category. Descriptive statistics were used to summarize the total sample and distribution of socio-demographic characteristics across study groups. We used Chi-square tests to compare the distribution of characteristics across study groups.

ROC Analysis: Receiver operating characteristic (ROC) curve analysis, plotting sensitivity vs (1-specificity), was used to examine OARS 12 item predictive validity compared to final case status. Scores examined for use as a cutoff for diagnostic prediction included the OARS-IV "Total Symptom Count" (TSC) and the "Total Impairment Mean" (TIM). ROC analysis using the total sample yielded a cutoff score range on each of these measures that optimized sensitivity and specificity. Area under the

curve (AUC) results were used to compare the accuracy yielded by the selected cutoff score in the total sample and across socio-demographic and developmental subgroups.

Sensitivity and Specificity: Within the total sample, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated for the range of selected cutoffs on each score type. Cutoff scores that provided the best balance between sensitivity and specificity were selected for further analysis in socio-demographic and developmental subgroups. Within each subgroup, the sensitivity, specificity, PPV and NPV generated by the selected cutoff scores was calculated. The threshold for acceptable sensitivity and specificity was set at 70% based on recommendations for screening measures (Glascoe 2005). If a cutoff failed to perform above 70% sensitivity and specificity in a subgroup, two scores below the cutoff were then examined. All analyses were conducted in SAS version 9.4.

3.3 Results

Participant Characteristics

Of 1,815 children ages 2-5 analyzed in this sample, 70% were determined to have final case status of ASD. In the Non-ASD group, 83% were determined to have a developmental delay. The sample was predominantly male, reflective of the higher prevalence ratio of ASD in males compared to females and the majority were white, non-Hispanic. Only 5% of the total sample had below average Mullen composite score, indicative of delays in one or more of the measured developmental domains. There were significant differences in the distribution of sex, race, Hispanic ethnicity, federal poverty level and maternal education between the ASD and non-ASD groups (Table 1). In the ASD group, 29% of the sample reported income and household size meeting

requirements below the FPL compared to 57% of the non-ASD group. The ASD group also had higher maternal education attainment compared to the non-ASD group.

Determining OARS Cutoff Scores

The results of the total sample ROC curves comparing TSC and TIM to final case status are shown in Table 2. TSC and TIM cutoff scores were selected for further examination based on location on the total sample ROC curve. Sensitivities, specificities, PPV and NPV for TSC scores ranging from 7-10 and TIM scores ranging from 0.25-0.40 were calculated from the total sample ROC curve. All scores assessed met the criteria of producing sensitivities and specificities > 70% in the total sample. Scores examined also produced little variation in calculated high PPVs, consistently above 93%, but a wide range of NPVs, ranging from 60% to 93%. Based on the balance of sensitivity and specificity, a TSC score of 9 and a TIM score of 0.3 were selected for use in subsequent subgroup analyses. At the selected cutoffs, both TIM and TSC scores had AUCs above 0.80 in the total and demographic subgroup samples (Table 3). Both score types performed with good accuracy (AUC >0.80) across sex, age, race, Hispanic ethnicity, poverty level, maternal education and developmental subgroups.

Table 1. Distribution of Participant Characteristics across Final Case Status

	ASD	Non-ASD	
Socio-demographic Characteristic	*N (%)*	*N (%)*	χ^2 *(df), p-value*
Sex			55.4 (1) <0.0001
Male	1051 (82%)	349(66%)	
Female	233 (18%)	182 (44%)	
Age			6.87 (3), 0.071
2 years	22 (2%)	8 (2%)	
3 years	156 (12%)	69 (13%)	
4 years	517 (40%)	180 (34%)	
5 years	589 (46%)	274 (52%)	
Race[1]			30.7 (2), <0.0001
White	681 (56%)	225 (47%)	
Black/African-American	283 (24%)	175 (37%)	
Other	236 (20%)	77 (16%)	
Hispanic[2]			5.58 (1), 0.018
Yes	233 (18%)	121 (23%)	
No	1038 (82%)	400 (77%)	
Federal Poverty Level (FPL)[3]			125.7(2) <0.0001
Below FPL	361 (29%)	279 (57%)	
At or above FPL	330 (27%)	108 (22%)	
Above 200% FPL	549 (44%)	104 (21%)	
Maternal Education[4]			117.5 (4)<0.0001
High school or Less	183 (15%)	179 (34%)	
Some College	411 (33%)	188 (36%)	
Bachelor's Degree	402 (31%)	98 (19%)	
Advanced Degree	270 (21%)	57 (11%)	
Mullen Early Learning Composite Score			0.256 (1), 0.613
Below Average	31 (2%)	15 (3%)	
At or Above Average	1253 (98%)	516 (97)	

1. Race, missing= 128
2. Hispanic, missing= 23
3. FPL, missing= 84
4. Education, missing= 27

Table 2. Performance of OARS in the Total Sample for ROC- Selected Range of Cutoffs

	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value PPV (95%CI)	Negative Predictive Value NPV (95%CI)
Total Symptom Count				
7	98 (96-99)	82 (79-86)	93 (92-94)	93 (90-94)
8	94 (93-96)	88 (85-90)	94 (93-96)	86 (84-89)
9	88 (87-90)	92 (90-95)	97 (95-98)	77 (73-80)
10	76 (73-78)	96 (94-98)	98 (97-99)	62 (59-65)
Total Impairment Mean				
0.25	96 (95-97)	85 (82-88)	94 (92-95)	90 (88-93)
0.30	92 (90-94)	90 (88-93)	96 (95-97)	83 (79-86)
0.35	85 (82-87)	94 (92-95)	97 (96-98)	72 (68-75)
0.40	73 (71-76)	96 (93-97)	98 (97-98)	60 (56-63)

Table 3. Calculated Area Under the Curve (AUC) for OARS across Socio-Demographic and Developmental Subgroups at Selected Cutoffs

	Total Symptom Count Cutoff=9	Total Impairment Mean Cutoff=0.3
	AUC (95% CI)	**AUC (95% CI)**
Total Sample	0.90 (0.88-0.92)	0.91 (0.89-0.93)
Sex		
Male	0.89 (0.88-0.92)	0.91 (0.89-0.93)
Female	0.92 (0.89-0.95)	0.92 (0.89-0.925
Age		
2 years	0.98 (0.93-1)	0.98 (0.92-1)
3 years	0.91 (0.87-0.95)	0.92 (0.88-0.96)
4 years	0.89 (0.86-0.92)	0.9 (0.86-0.92)
5 years	0.91 (0.88-93)	0.96 (0.90-94)
Race		
White	0.89 (0.87-0.92)	0.90 (0.88-0.93)
Black	0.93 (0.91-0.95)	0.94 (0.92-0.96)
Other	0.91 (0.87-0.95)	0.90 (0.87-0.95)
Hispanic		
Yes	0.86 (0.81-0.88)	0.85(0.82-0.91)
No	0.91 (0.90-0.93)	0.93 (0.92-0.95)
Federal Poverty Level (FPL)		
Below FPL	0.90 (0.88-0.93)	0.92 (0.90-0.94)
At or above FPL	0.92 (0.89-0.95)	0.91 (0.88-0.95)
Above 200% FPL	0.88 (0.85-0.92)	0.91 (0.87-0.94)
Maternal Education		
High school or Less	0.89 (0.86-0.92)	0.88 (0.85-0.92)
Some College	0.92 (0.89-0.94)	0.93 (0.91-0.95)
Bachelor's Degree	0.93 (0.91-0.96)	0.94 (0.92-0.97)
Advanced Degree	0.87 (0.83-0.92)	0.89 (0.84-0.93)
Mullen Early Learning Composite Score		
Below Average	0.85 (0.74-0.97)	0.86 (0.75-0.98)
At or Above Average	0.90 (0.87-0.92)	0.89(0.88-0.91)

Performance across Subgroups

Table 4 shows the ranges of sensitivity, specificity, PPV and NPV calculated for the selected TSC and TIM scores across socio-demographic and developmental subgroups. TIM had a slightly higher sensitivity, but lower specificity compared to TSC. Both scores performed above the 70% criteria for sensitivity and specificity across all socio-demographic subgroups. Across racial, age, and poverty categories, minor variation was found in sensitivity of either tool, while specificity had greater variation within these subgroups. TSC had higher sensitivity (91%) and specificity (83%) within the Hispanic group compared to TIM sensitivity (84%) and specificity (75%). Comparing performance across maternal education levels, the lowest sensitivities and specificities were found in the advanced degree subgroup. With regards to PPV and NPV, TSC and TIM had greater variation in performance with neither score type consistently outperforming the other. Lower NPVs (<70%) were found in high income and advanced degree subgroups.

Within the MSEL-derived "below average" developmental subgroup, both TSC and TIM performed with sensitivities and specificities with wide confidence intervals which dipped below the 70% threshold. Alternate cutoff scores were examined to improve performance in the below average development group (Table 5). At lower cutoff scores, but TSC and TIM derived sensitivity and specificity point estimates improved. However, the small sample of this group contributed to wide confidence intervals, which contained estimates below the acceptable threshold.

Table 4. Table of Sensitivities and Specificities by Socio-Demographic and Developmental Characteristics

	Total Symptom Count				Total Impairment Mean			
Cutoff Score	9				0.3			
	Sensitivity	**Specificity**	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
Total Sample	88 (87-90)	92 (90-95)	97 (95-98)	77 (73-80)	92 (90-94)	90 (88-93)	96 (95-97)	83 (79-86)
Sex								
Male	88 (86-90)	92 (89-95)	97 (96-98)	72 (67-76)	92 (91-94)	89(86-93)	96 (95-98)	79 (75-83)
Female	91 (86-94)	92 (89-97)	94 (91-97)	88 (84-93)	92 (88-95)	92 (87-96)	93 (90-96)	90 (86-95)
Age								
2 years	95 (87-100)	100	100	88 (68-100)	95 (86-100)	100	100	89 (73-100)
3 years	88 (83-93)	94 (89-99)	97 (94-100)	77 (68-86)	94 (91-98)	90 (89-96)	95 (92-98)	87 (80-95)
4 years	89 (86-91)	89 (85-94)	96 (94-97)	73 (76-79)	91 (89-94)	87 (81-92)	95 (93-97)	78 (73-84)
5 years	88 (86-91)	93 (90-96)	96 (95-98)	79 (74-83)	92 (90-94)	92 (89-95)	96 (95-98)	84 (80-88)
Race								
White	88 (85-90)	91(88-95)	97 (95-98)	71 (66-76)	92 (90-94)	88 (84-93)	96 (95-98)	79 (74-84)
Black	89 (85-93)	97 (94-99)	98 (96-99)	85 (79-90)	92 (89-95)	97 (93-99)	98 (96-99)	88 (84-93)
Other	90 (86-94)	92 (86-98)	97 (95-99)	74 (65-83)	92 (89-96)	90 (83-96)	97 (94-99)	78 (70-87)
Hispanic								
Yes	91 (87-95)	83 (76-89)	91 (87-95)	83 (76-89)	84 (91-97)	75 (70-81)	88 (83-92)	87 (81-94)
No	88 (86-90)	95 (93-97)	97 (96-99)	75 (71-79)	95 (92-97)	92 (90-93)	82 (78-85)	98 (97-99)
Federal Poverty Level (FPL)								
Below FPL	88 (84-91)	93 (90-96)	94 (92-97)	85 (81-90)	92 (89-96)	91 (88-94)	89 (85-92)	94 (92-96)
At or above FPL	90 (86-93)	94 (90-99)	98 (96-99)	76 (68-82)	93 (90-96)	90 (84-96)	96 (94-98)	81 (73-87)
Above 200% FPL	88 (85-90)	89 (83-95)	98 (96-99)	59 (51-66)	92 (90-95)	88 (82-95)	97 (96-99)	69 (61-77)
Maternal Education								
High school or Less	89 (84-93)	90 (85-94)	90 (86-94)	88 (83-93)	90 (85-94)	87 (81-92)	87 (82-92)	89 (84-94)
Some College	90 (87-93)	94 (90-97)	97 (95-98)	81 (75-86)	94 (92-96)	92 (88-96)	96 (94-98)	87 (82-92)
Bachelor's Degree	91 (88-94)	94 (90-99)	99 (97-99)	73 (64-80)	95 (93-97)	94 (89-99)	98 (97-99)	81 (74-89)
Advanced Degree	83 (78-88)	91 (84-99)	98 (96-99)	53 (43-65)	88 (84-92)	89 (82-97)	97 (96-99)	61 (51-72)
MSEL								
Below average	71 (54-86)*	93 (80-100)	96 (87-100)	61 (40-80)	74 (58-89)*	80 (59-100)	88 (76-100)	60 (38-82)
At or above average	89 (87-91)	92 (90-95)	96 (95-97)	77 (74-91)	93(91-94)	91 (88-93)	96 (95-97)	84 80-87)

*Sensitivity or specificity falls at or below the recommended minimum 70% threshold

Table 5. Sensitivity and Specificity of Alternate Cutoffs in the Mullen Developmental Subgroup

	Total Symptom Count				**Total Impairment Mean**			
Cutoff Score	7		8		0.2		0.25	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Early Learning Composite Score								
Below average	90 (79-100)	80 (59-100)*	84 (70-96)	80 (60-100)*	94 (84-100)	80 (72-100)	94 (85-100)	80 (60-100)*
At or above average	98 (96-99)	82 (79-85)	95(93-96)	88 (84-91)	93 (91-94)	91 (88-93)	96 (93-97)	86 (82-89)

*Sensitivity or specificity falls at or below the recommended minimum 70% threshold

3.4 Discussion

Our results suggest that the OARS 12 item may be suitable tool for use in identifying ASD among high-risk children ages 2-5 years. Generally, the sensitivity and specificity of the tool were well above the recommended threshold for diagnostic accuracy across all socio-demographic subgroups. Generally, the "Total Symptom Count" had higher specificity across socio-demographic subgroups, while "Total Impairment Meant" had higher sensitivity. Researchers and clinicians who use OARS 12 will need to consider which psychometric feature they wish to prioritize based on how the measure is being used.

Our analyses did not show a decrease in diagnostic accuracy for lower SES or racial/ethnic minority groups as has been reported by studies of similar measures (Moody et al, 2017; Scarpa et al. 2013) With the exception of the advanced degree subgroup, these results follow the reported trend of decreasing screening accuracy with lower maternal educational attainment (Moody et al. 2017). The lower sensitivity and specificity of the advanced degree subgroup compared to lower maternal education sub-groups was an unexpected result and could not be explained by a higher percentage of children with below average developmental scores, as this group also had reduced OARS tool performance. Lower OARS 12-item sensitivity and specificity among children scoring "below average" across MSEL developmental domains suggest that alternative thresholds or subscales should be explored in larger population samples for children at lower developmental levels. Prior work in the SEED study has shown decreased specificity of other measures such as the Social Responsiveness Scale and the Social Communication Questionnaire among children with below average developmental level as measured by the MSEL (Moody et al. 2017).

This study has several limitations. Primarily, the OARS instrument was completed after ADOS and ADI-R assessments and often by the same clinician that administered these assessments. In this way, OARS 12-item responses were informed by child performance on ADOS and ADI-R items as well as caregiver discussions and overall evaluation visit. This likely resulted in higher OARS accuracy than would otherwise have occurred if the tool had been administered at a different time in the study visit. Thus, our results serve as a proof-of-concept in an ideal setting and must be tested in a true high risk screening scenario. The result remain valuable as this validity had not previously been documented, even in this type of setting.

Limited information was available on the fidelity of evaluators to the OARS item language and author-recommended administration methods across research sites. Also, the OARS tool was only administered to participants undergoing ASD evaluation having initially screened positive for ASD risk, to the exclusion of the majority of the general population-based controls. This may have effectively reduced the available true negatives in the sample. Finally, the underrepresentation of non-white, Hispanic, and below average development subgroups in the sample limits the generalizability of these findings.

These analyses suggest that the OARS 12-item has psychometric properties similar to other recognized methods of ASD measurement that allow for the identification of ASD among high risk children age 2-5. While the OARS was not administered as a screening tool in this study, efforts could be taken to adapt these items into a more structured tool for use as a level-2 screener. As an observation-based, no-cost tool, the OARS potentially offers a complimentary measurement approach to existing

identification pathways. In lower resource settings where long wait times and high volume of children requiring evaluation result in delays in accessing early intervention services, an adaptation of the OARS items could be implemented to provide an intermediate measure of ASD between initial pediatric screening and formal evaluation by a team of specialists. More research is needed on OARS performance in the types of community and clinical settings in which it could serve as an additional no-cost instrument to assist in the multi-step identification of ASD.

3.5 References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (5th ed.). Arlington, VA: American Psychiatric Association.

Austin, J., Manning-Courtney, P., Johnson, M. L., Weber, R., Johnson, H., Murray, D., et al. (2016). Improving access to care at autism treatment centers: A system analysis approach. *Pediatrics, 137 Suppl 2*, S149-57.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C.), 67*(6), 1-23.

Bisgaier, J., Levinson, D., Cutts, D. B., & Rhodes, K. V. (2011). Access to autism evaluation appointments with developmental-behavioral and neurodevelopmental subspecialists. *Archives of Pediatrics & Adolescent Medicine, 165*(7), 673-674.

Centers for Disease Control and Prevention (CDC). (2015). *Autism spectrum disorder: Screening and diagnosis for health care providers.* Retrieved August 8, 2016, from http://www.cdc.gov/ncbddd/autism/hcp-screening.html

Centers for Disease Control and Prevention (CDC). (2018). *Study to explore early development (SEED).* Retrieved 8/10, 2018, from https://www.cdc.gov/ncbddd/autism/seed.html

de Bildt, A., Sytema, S., Zander, E., Bölte, S., Sturm, H., Yirmiya, N., et al. (2015). Autism diagnostic interview-revised (ADI-R) algorithms for toddlers and young

preschoolers: Application in a non-US sample of 1,104 children. *Journal of Autism and Developmental Disorders, 45*(7), 2076-2091.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(3), 173-179.

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics, 63*(5), 851-859.

Hagan, J. F., Shaw, J. S., & Duncan, P. M. (2007). *Bright futures: Guidelines for health supervision of infants, children, and adolescents* (4th ed.). Elk Grove Village, IL.: Am Acad Pediatrics.

Jo, H., Schieve, L. A., Rice, C. E., Yeargin-Allsopp, M., Tian, L. H., Blumberg, S. J., ... & Boyle, C. A. (2015). Age at autism spectrum disorder (ASD) diagnosis by race, ethnicity, and primary household language among children with special health care needs, United States, 2009–2010. *Maternal and Child Health Journal*, *19*(8), 1687-1697.

Koegel, L. K., Koegel, R. L., Nefdt, N., Fredeen, R., Klein, E. F., & Bruinsma, Y. E. (2005). First STEP: A model for the early identification of children with autism spectrum disorders. *Journal of Positive Behavior Interventions*, *7*(4), 247-252.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685.

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal, 14*(3), 234-252.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205-223.

Moody, E. J., Reyes, N., Ledbetter, C., Wiggins, L., DiGuiseppi, C., Alexander, A., ... & Rosenberg, S. A. (2017). Screening for autism with the SRS and SCQ: Variations across demographic, developmental and behavioral factors in preschool children. *Journal of Autism and Developmental Disorders*, *47*(11), 3550-3561.

Mullen, E. M. (1995). *Mullen scales of early learning* AGS Circle Pines, MN.

Olsson, N. C., & Bölte, S. (2014). Brief report: "quick and (not so) dirty" assessment of change in autism: Cross-cultural reliability of the developmental disabilities CGAS and the OSU autism CGI. *Journal of Autism and Developmental Disorders, 44*(7), 1773-1778.

Olsson, N. C., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., et al. (2017). Social skills training for children and adolescents with autism spectrum disorder: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(7), 585-592.

OSU Research Unit on Pediatric Psychopharmacology (OSU RUPP). (2005). *OSU autism rating scale-DSM 4 (OARS-4)*. Columbus, Ohio: OSU Research Unit on Pediatric Psychopharmacology.

Robins, D. L., Adamson, L. B., Barton, M., Connell, J. E., Dumont-Mathieu, T., Dworkin, P. H.,… & Newschaffer, C. (2016). Universal autism screening for toddlers: Recommendations at odds. *Journal of autism and developmental disorders*, *46*(5), 1880-1882.

Rutter M, Bailey A, Lord C. Social Communication Questionnaire (SCQ). Los Angeles: Western Psychological Services; 2003.

Scarpa, A., Reyes, N. M., Patriquin, M. A., Lorenzi, J., Hassenfeldt, T. A., Desai, V. J., & Kerkering, K. W. (2013). The modified checklist for autism in toddlers: Reliability in a diverse rural American sample. *Journal of Autism and Developmental Disorders*, *43*(10), 2269-2279.

Schendel, D. E., DiGuiseppi, C., Croen, L. A., Fallin, M. D., Reed, P. L., Schieve, L. A., et al. (2012). The study to explore early development (SEED): A multisite epidemiologic study of autism by the centers for autism and developmental disabilities research and epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders, 42*(10), 2121-2140.

Schulz E. (2016). *Executive summary: Capacity building for ASD and DD: Community Based autism liaison and training (CoBALT) project in Arkansas*

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of autism and developmental disorders*, *34*(6), 691-701.

Stone, W. L. (2013). Screening tool for autism in two-year-olds (STAT). *Encyclopedia of autism spectrum disorders*, 2684-2688.

U.S Department of Health and Human Services. (2018). *HHS 2012 poverty guidelines.* Retrieved 7/1, 2018, from https://aspe.hhs.gov/2012-hhs-poverty-guidelines

Warren, Z., Stone, W., & Humberd, Q. (2009). A training model for the diagnosis of autism in community pediatric practice. *Journal of Developmental & Behavioral Pediatrics, 30*(5), 442-446.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-11.

Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., DiGuiseppi, C., et al. (2015). Autism spectrum disorder symptoms among children enrolled in the study to explore early development (SEED). *Journal of Autism and Developmental Disorders, 45*(10), 3183-3194.

Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., et al. (2015). Using standardized diagnostic instruments to classify children with autism in the study to explore early development. *Journal of Autism and Developmental Disorders, 45*(5), 1271-1280.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., et al. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136 Suppl 1*, S41-59.

CHAPTER 4: CONSTRUCT VALIDITY OF THE OHIO STATE AUTISM RATING SCALE (OARS 12-ITEM) IN THE STUDY TO EXPLORE EARLY DEVELOPMENT

4.1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as a restrictive and repetitive pattern of behaviors or interests (APA, 2013). Detection of early signs and symptoms of ASD before the age of three is key to providing appropriate interventions to reduce ASD-related impairment during this critical early developmental window (Zwaigenbaum et al. 2015; Warren et al. 2011). The early identification pathway for ASD in the U.S. relies on detection of early ASD signs and symptoms in pediatric primary care through the use of routine ASD-specific and broader developmental screening instruments (Zwaigenbaum et al. 2015; Hagan et al. 2017). Once an initial threshold of concern is met using one of the commonly used "level-one" screening tools such as the Modified Checklist for Autism in Toddlers (M-CHAT), referral is made for diagnostic evaluation from a developmental specialist (Robins 2016; Zwaigenbaum et al. 2015). In communities with already overburdened clinical infrastructures, the limited supply of developmental subspecialists with training on ASD evaluation, the high cost of per-use licensed screening and assessment materials, and the recent increase in the number of children needing early ASD evaluation and referral can lead to delays in providing appropriate ASD identification. This can disproportionately affect families who already have a greater burden of health disparities and economic disadvantage (Gordon-Lipkin et al. 2017; Warren et al. 2009; Schulz, 2016).

One potential ASD measurement tool that might complement existing screening and evaluation procedures in low-resource settings is the no-cost Ohio State Autism Rating Scale (OARS). Published for public use in 2005 and based on DSM-IV ASD domains, the 12-item OARS is an interactive measurement tool that uses clinician observation as well as a semi-structured interview with the primary care giver (OSU RPP, 2005). Unlike many of the other screening and assessment tools commonly used for ASD, the OARS 12-item does not have a per-use license fee and is free to use for clinicians and researchers. It is also relatively brief; the OARS authors report an average 30 minute administration time. This brief scale measures the three DSM-IV domain criteria through a series of 12 questions rated on an ordinal scale of "Never or Rarely; not a problem" to "Very Often; a severe problem". The OARS obtains measures of ASD severity through a calculation of two scores: "Total Symptom Count", derived by counting all items with greater than a zero rating, and "Impairment Mean", derived by averaging the severity scores of each item across domains. While not as structured as other observation-based tools such as the Childhood Autism Rating Scale (CARS), the OARS 12-item offers a no-cost option to be added alongside existing screening and assessment tools to offset the higher volume of need and reduce delays to obtaining sufficient evaluation to access early intervention.

Prior work in the Study to Explore Early Development (SEED 1&2) has shown the OARS 12-item to be valid compared to gold-standard tools for ASD assessment and thus may be a useful a freely available, brief assessment tool that could be used when needed in low-resourced settings (Chapter 3). In that analysis, the OARS 12-item demonstrated acceptably high (>0.80) levels of sensitivity and specificity across gender,

racial/ethnic, maternal education, and income subgroups (Chapter 3). However, among children with below average development based on the Mullen Scales of Early Learning, the OARS performed with reduced sensitivity (0.71) and specificity (0.74).

Beyond criterion validity, there is a need to assess how the 12 items of the OARS relate to and measure the underlying ASD constructs of social and verbal communication and repetitive behaviors, compared to more expensive gold standard tools. It is important to understand whether the OARS instrument's ability to capture underlying ASD constructs varies by socio-demographic or developmental subgroups, which may in turn influence the performance of this tool within diverse populations.

The SEED study has available OARS 12-item information alongside gold standard tools for children with suspected ASD , which makes examination of item performance, comparison to gold standard tools, and comparison across subgroups possible. However, the approach has limitations given that clinicians were aware of the results of gold-standard ADOS and ADI-R tools when completing OARS items, as has been discussed in previous work (Chapter 3). This analysis of OARS 12-item construct validity seeks to add further proof-of-principle that the OARS may be a useful free tool by examining how the OARS measures specific underlying ASD constructs in this high-risk group, and whether this measurement varies by socio-demographic or developmental subgroup.

4.2 Methods

Participants

SEED is a multi-site case-control study conducted in multiple phases at research institutions across the U.S since 2008 (Schendel et al. 2012; Centers for Disease Control

and Prevention, 2018). SEED was established to gather information about ASD behavioral phenotype and associated medical, developmental and behavioral conditions as well as to examine possible environmental and genetic risk factors for autism among children age 2-5 (Schendel et al. 2012; Wiggins et al, 2015). Participants were recruited from sites in California, Colorado, Georgia, Maryland, North Carolina, and Pennsylvania between the ages of 30-64 months. SEED phase 1 recruited children from 2007-2011 and phase II from 2012-2016.

Typically-developing children were recruited from the general population based on a random sampling of state vital records. Children with potential developmental delay (DD) or ASD were invited to participate after being identified from educational and healthcare providers that serve children with DD. Children were excluded from participation in the SEED study if they were born outside the designated study catchment areas, had mobility, vision or hearing impairments that prevented full engagement in the assessments, or did not meet spoken language requirements. Participants with a prior ASD diagnosis or who scored above an 11 (Wiggins et al 2015) on the Social Communication Questionnaire (Rutter, Bailey and Lord, 2003) in the first phone interview went on to receive an Autism evaluation in addition to a developmental assessment and collection of bio-samples. Participants not meeting these requirements received only developmental assessment and bio-sample collection. During the autism evaluation component, autism specific parent interviews and child observation measures were performed. Participants without complete data on the OARS tool or who did not receive a final case determination were excluded from this analysis.

Measures

Autism Diagnostic Observation Schedule (ADOS)

The ADOS is a standardized diagnostic measure that involves direct interaction and scoring of child's behavior, taking into account developmental level and age of the child. Modules 1, 2, and 3 were performed based on child age and cognitive level. ADOS items cover a range of social communication items such as "response to join attention" and "facial expression" as well as items measuring restricted and repetitive patterns of interest or behavior such as "unusual sensory interests" (Gotham et al. 2007). Prior analyses of the ADOS in the SEED 1 sample found sensitivity and specificity of 0.91 and 0.62 respectively (Wiggins et al. 2015b). In this analysis, case status was determined by scores ranging from 7-11 depending on functional ability (Wiggins et al. 2015b).

Autism Diagnostic Interview-Revised

The ADI-R is a semi-structured parent/caregiver interview that includes 93 questions about ASD characteristics across the three DSM-IV-TR domains of language/communication, reciprocal social interaction and restricted, repetitive and stereotyped behaviors and interests (Lord et al. 1994) All ADI-R items are coded for past and current behavior. Scoring is coded using different algorithms based on age and can be administered to caregivers of children as young as 12 months (de Bildt et al. 2015). The ADI-R has a reported sensitivity and specificity of over 0.90 (Lord et al. 1993). In this analysis, autism cutoff score was determined by 10 on social deficits, 7-8 on communication deficits, and 3 on behavioral deficits (Wiggins et al. 2015).

Mullen Scales of Early Learning (MSEL)

The MSEL is a developmental assessment comprised of five scales: gross motor, fine motor, visual reception, expressive language and receptive language (Mullen, 1995). Completion of a variety of progressively more difficult tasks in the 5 domains produces an early learning composite score as well as 4 domain sub-scores. Corresponding t-scores are selected based on sub scores and child age at evaluation. In this analysis, a t-score below 40 on any domain sub-score or below 85 on the early learning composite score was considered "below average".

Ohio State Autism Rating Scale (OARS 12 item)

The OARS 12 item is a semi-structured interview administered scale for clinicians and trained researchers with three DSM-IV based subscales for social impairment, communication impairment and restricted, repetitive and stereotyped behavior. The OARS has 12 questions rated on an ordinal scale of "0: Never or Rarely; not a problem" to "3: Very Often; a severe problem". The OARS obtains measures of ASD severity through a calculation of two scores: "Impairment Mean" and a "Total Symptom Count". Symptom count for each domain and for total symptom count is calculated by adding the number of symptoms that received a score of greater than "0: Rarely or Never". The impairment mean for each domain and total impairment mean are calculated by adding the number of points 0-3 gained for severity of each item then dividing by the number of points possible. The denominators for verbal children are 12 for social interaction, 12 for communication, 12 for restricted patterns. The denominators for nonverbal children are 12 for social interaction, 9 for communication (excluding item B2), and 12 for restricted patterns (OSU RPP, 2005).

Final Case Status

Final case determination of ASD and DD status was made using a previously published algorithm (Wiggins et al 2015) based primarily on the ADI-R and ADOS. If the child met SEED ASD criteria but had a mental age less than 24 months, the OARS degree of certainty (DOC) rating (added for SEED) was used to determine case status (Wiggins et al. 2015). Specifically, if clinicians were certain the child had ASD (i.e. DOC rating of 4 or 5) the child was classified as an ASD case and if clinicians were uncertain the child had ASD (i.e., DOC rating of 1, 2, or 3) the child was classified as "indeterminate." Indeterminate cases were subsequently dropped from analyses because ASD could not be confirmed, or could not be distinguished from intellectual disability.

Maternal Interview

Mothers reported socio-demographic factors via SEED questionnaires. Of particular interest for this study were maternal and paternal race/ethnicity, maternal educational attainment, maternal employment patterns during pregnancy and estimated household number of residents and income during pregnancy and at time of study enrollment. Income level relative to the Federal Poverty Level (FPL) was calculated using guidelines and thresholds for 2012 and accounted for number of persons in household as well as income for the year prior to the child's evaluation (HHS, "Poverty Guidelines" 2012).

Study Sample

For this analysis, we included all children enrolled in SEED 1 or SEED 2 who received a full diagnostic evaluation for ASD, regardless of final case status designation.

This includes children with past ASD diagnosis, children who screened positive on the SCQ, and children who raised clinician concern at the developmental evaluation. Children with incomplete information on the OARS were excluded.

Statistical Analysis

Multi-trait multimethod (MTMM) correlation matrices are an approach used to assess construct validity of a set of measures. A MTMM table produces a set of correlations arranged to facilitate interpretation of *convergent validity*, the evidence of similarity between measures of theoretically related constructs, and *discriminant validity*, the absence of correlation between measures of theoretically unrelated constructs (Campbell & Fiske, 1959; DeVellis 2012). A MTMM matrix was used to assess correlation between OARS impairment mean score, ADOS, and ADI-R on social communication, interaction and restricted or repetitive behaviors or interest (RRBI) domains. Impairment mean score was selected for use in this analysis over Total Symptom Count because of its greater sensitivity across subgroups in previous SEED work (Chapter 3). Correlation between ADOS & ADI-R domain scores was also assessed. Correlations were considered very strong if greater than 0.80, strong if greater than 0.60 and moderate if greater than 0.40 (DeVellis 2012).

Exploratory Factor Analyses (EFA) were conducted to explore the underlying structure of how the OARS 12 items map onto and represent ASD traits, or "factors". Using Mplus version 5.2, the 12 items of the OARS with categorical outcomes were submitted to EFA with maximum likelihood estimation. The principal factor method was used to extract factors and the relative improvement in the fit of each additional factor was assessed. A scree test suggested two meaningful factors. Only these factors were

retained for oblique promax rotation, assuming correlated factors. Items were considered cross-loaded if they had loadings greater that 0.30 on one or more factor.

Using the 2-factor structure gained from EFA results, a series of confirmatory factor analysis (CFA) models were fit to measure the equality of the factor structure across socio-demographic and developmental subgroups. Items that cross-loaded during EFA were constrained to load on the factor with the highest loading. Four CFA models ranging from least to most restrictive were fit to assess measurement invariance: configural, metric, scalar, strict. The model assessing *configural invariance* (also called pattern invariance) is least restrictive and evaluates whether constructs have the same pattern of free and fixed loading across subgroups, without applying any equality constraints between the groups (Putnick & Borstein, 2016). The model assessing *metric invariance* constrains factor loadings to be equivalent between two of more subgroups and evaluates whether each item contributes to the constructs to the same degree across subgroups. The model assessing *scalar invariance* constrains item intercepts to be equivalent between two or more subgroups and evaluates whether the constructs are measured on the same scale across subgroups using mean differences. The final and most stringent model assessed *strict invariance* and constrained factor loadings, item intercepts and residual variance to be equal across subgroups.

Comparisons between measurement invariance models were done using an equivalence testing approach. Chi-square tests of model fit between each new measurement invariance model were performed to generate p-values. However, the chi-square test's sensitivity to sample size limits its usefulness for evaluation of measurement invariance. Rather, multiple fit indices were prioritized above chi-square and chi-square

difference tests for model fit evaluation to address potential model misspecification errors (Jian, Mai & Yuan, 2017). The four measurement invariance models were compared using three fit indices: root-mean-squares-error of approximation (RMSEA), comparative fit index (CFI), and the Tucker-Lewis index (TLI) (Jian, Mai & Yuan, 2017; Putnick & Borstein, 2016). A drop in CFI or TLI greater than or equal to 0.01 or an increase in RMSEA greater than or equal to 0.01 implies measurement nonequivalence between selected subgroups, therefore a change in either was considered sufficient to not move to the next modeling step. (Jiang et al, 2017; Chen, 2007; Cheung and Rensvold, 2002).

4.3 Results

The analytic sample included 1,933 children age 2-5 years with the majority of children over age 3 (Table 1). The sample was predominantly male, reflective of the higher prevalence of ASD among males. Race/ethnicity was analyzed as three categories, with White being the largest racial/ethnic group represented. Regardless of race, a total of 35% of the sample reported Hispanic ethnicity and there was not a significant difference in this distribution across ASD and Non-ASD case groups. Income and number of household members were used to calculate federal poverty level relative placement, with 35% of the total sample reporting income and household size below the FPL threshold. Participants in the Non-ASD group reported lower FPL status compared to the ASD group. The smallest subgroup included in this analysis included children who scored below average on the Mullen-derived Early Learning Composite Score for development.

Table 1: Distribution of SEED Participant Demographics

Socio-demographic Characteristic	**ASD**	**Non-ASD**	
	N (%)	*N (%)*	*χ^2, (df), p-value*
Sex			46.2 (1), <0.001
Male	1042 (81.6)	445 (67.8)	
Female	235 (18.4)	211(32.2)	
Age			4.96 (3), 0.175
2 years	25 (2)	13 (2)	
3 years	166 (13)	86 (13.2)	
4 years	521 (40.8)	235 (35.8)	
5 years	565 (44.2)	322 (49)	
Race[1]			32.6 (2), <0.001
White	685 (56.8)	284 (41.1)	
Black/African-American	279 (23.1)	216 (35.8)	
Other	242 (20.1)	103 (17.1)	
Hispanic[2]			0.278 (1), 0.598
Yes	224 (17.7)	120 (18.7)	
No	1039 (82.3)	521 (81.3)	
Federal Poverty[3] **Level (FPL)**			107.6 (2), <0.001
Below FPL	331 (26.8)	308 (50.4)	
At or above FPL	338 (27.4)	142 (23.2)	
Above 200% FPL	565 (45.8)	161 (26.4)	
Maternal Education[4]			91.6 (4) <0.0001
High school or Less	163 (13)	185 (28.2)	
Some College	413 (32.5)	232 (36.4)	
Bachelor's Degree	402 (32)	138 (21.2)	
Advanced Degree	282 (22.5)	90 (14.2)	
Mullen Early Learning Composite Score			2.05 (1), 0.153
Below average	32 (2.5)	24 (3.7)	
At or Above average	1245 (97.5)	632 (96.3)	

1. Race: Missing = 124
2. Hispanic: Missing = 29
3. FPL: Missing = 88
4. Maternal Education: Missing = 28

The MTMM correlation matrix shows correlations between measures of similar ASD domains, as measured by the different modalities of OARS, ADOS and ADI-R (Table 2). Unlike OARS and ADI-R, the ADOS only has two domains available for comparison, a composite "social affect" domain that combines social interaction and communication and a restricted and repetitive pattern of behaviors (RRB) domain comparable to that of OARS and ADI-R. Very strong correlation (0.81) was found between OARS social

interaction mean and ADOS social affect. Strong correlations (> 0.60) were found between OARS and ADOS or ADI-R for several domains including social interaction/deficits. Moderate correlation was found between OARS and ADI-R measurements of social interaction, communication and RRB domains. There was weak correlation between ADOS and ADI-R RBB measures.

The results of the EFA are shown by the pattern and structure matrix coefficients in tables 3 and 4. Only loadings of greater than 0.30 were reported. In Table 3, item 1 assessing impairment in the use of non-verbal behaviors did not load strongly on either factor. Items 2-6 and 9 uniquely loaded on Factor 1, while items 8, 10 and 11 uniquely loaded on Factor 2. Items 7 and 12 cross-loaded on both factors, indicating a contribution of both factors in that item's measurement. Factors 1 and 2 were assigned ASD domain names based on the items which loaded to each factor. Moderate to strong correlations were found across all items except item 1, with several items strongly correlated to both factors (Table 4).

Table 2. Multi-Trait Multi-Method Correlation Matrix Comparing OARS, ADOS, and ADI-R Domains

	OARS Social Interaction Mean	OARS Social Comm. Mean	OARS Restrictive/ Repetitive Behaviors Mean	ADOS Social Affect Total Score	ADOS Restrictive and Repetitive Behaviors Total Score	ADIR Social Deficits Total Score	ADIR Verbal/Non-Verbal Comm. Deficits Total Score	ADIR Behavioral Deficits Total Score
OARS Social Interaction Mean	1.00							
OARS Social Communication Mean	0.81	1.00						
OARS Restrictive/Repetitive Behaviors Mean	0.65	0.62	1.00					
ADOS Social Affect Total Score	0.81	0.72	0.56	1.00				
ADOS Restrictive and Repetitive Behaviors Total Score	0.52	0.53	0.61	0.54	1.00			
ADIR Social Deficits Total Score	0.72	0.63	0.57	0.59	0.43	1.00		
ADIR Verbal Communication Deficits Total Score	0.50	0.46	0.43	0.47	0.35	0.67	1.00	
ADIR Behavioral Deficits Total Score	0.37	0.31	0.59	0.33	0.32	0.49	0.54	1.00000

Light shading indicates correlation between OARS & ADOS /ADI-R on similar domains

Dark shading indicates correlation between gold-standard ADOS and ADI-R themselves for similar domains

Table 3. Rotated Exploratory Factor Analysis Factor Pattern Coefficients

	Factor Loadings	
OARS Item	*Social Communication & Interaction*	*Repetitive or Restricted Patterns of Behavior or Interest*
Impairment in use of nonverbal behaviors to regulate social interaction	--	--
Impairment in peer relations	0.78	--
Impairment in spontaneous seeking to share enjoyment, interests, or achievements with other people	0.56	--
Impairment in social or emotional reciprocity	0.83	--
Does not attempt to speak or communicate; if nonverbal, fails to use gesture or mime to communicate	0.77	--
If adequate speech: impairment in the ability to initiate or sustain a conversation.	0.78	--
Stereotyped and repetitive use of language or sounds or idiosyncratic language*	0.49	0.40
Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level	--	0.78
Excessive preoccupation with or stereotyped, restricted patterns of interest that are abnormal/odd either in intensity or focus	0.68	--
Inflexible adherence to specific, nonfunctional routines or rituals	--	0.63
Stereotyped, repetitive motor mannerisms	--	0.57
Preoccupation with parts of objects*	0.38	0.34

*Item cross-loads on more than 1 factor
Only factor loadings > 0.30 reported

Table 4. Rotated Exploratory Factor Analysis Factor Structure Coefficients

	Factor Structure	
OARS Item	*Social Communication & Interaction*	*Repetitive/ Restricted Patterns of Behavior/ Interest*
Impairment in use of nonverbal behaviors to regulate social interaction	0.13	0.09
Impairment in peer relations	0.87	0.66
Impairment in spontaneous seeking to share enjoyment, interests, or achievements with other people	0.73	0.64
Impairment in social or emotional reciprocity	0.88	0.63
Does not attempt to speak or communicate; if nonverbal, fails to use gesture or mime to communicate	0.90	0.71
If adequate speech: impairment in the ability to initiate or sustain a conversation.	0.79	0.55
Stereotyped and repetitive use of language or sounds or idiosyncratic language	0.76	0.73
Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level	0.49	0.75
Excessive preoccupation with or stereotyped, restricted patterns of interest that are abnormal/odd either in intensity or focus	0.81	0.66
Inflexible adherence to specific, nonfunctional routines or rituals	0.54	0.71
Stereotyped, repetitive motor mannerisms	0.33	0.53
Preoccupation with parts of objects	0.60	0.59

A CFA model was fit using the previously derived 2-factor structure, where item 1 was excluded and items 7 and 12 were constrained to load on the factor with the highest loading. Table 5 shows RMSEA, CFI and TLI fit indices for each of the measurement invariance models as well as change in the indices for each subsequent model. The strictest form of measurement invariance was found for the measurement model across all subgroups except FPL. The change in CFI between the scalar and strict models did not support measurement equivalence between FPL subgroups, indicating that the additional constraint of equal residual variances of the model across subgroups could not be met.

Table 5. Fit Indices for Testing Measurement Invariance in CFA Models

	RMSEA	Change in RMSEA	CFI	Change in CFI	TLI	Change in TLI	Invariance Rule Accepted?
Age							
Configural[1]	0.084	--	0.980	--	0.975	--	
Metric[2]	0.080	-0.004	0.980	0	0.978	+0.003	Yes
Scalar[3]	0.066	-0.014	0.983	+0.003	0.985	+0.007	Yes
Strict[4]	0.061	-0.005	0.985	+0.002	0.987	+0.002	Yes
Sex							
Configural	0.086	--	0.978	--	0.972	--	
Metric	0.083	-0.003	0.977	-0.001	0.974	+0.002	Yes
Scalar	0.070	-0.013	0.981	+0.004	0.982	+0.008	Yes
Strict	0.071	+0.001	0.980	-0.001	0.981	+0.001	Yes
Race							
Configural	0.088	--	0.979	--	0.974	--	
Metric	0.083	-0.005	0.979	0	0.977	+0.003	Yes
Scalar	0.068	-0.015	0.983	+0.004	0.985	+0.008	Yes
Strict	0.070	+0.002	0.981	-0.002	0.984	-0.001	Yes
Hispanic							
Configural	0.084	--	0.981	--	0.976	--	
Metric	0.079	-0.005	0.981	0	0.978	+0.002	Yes
Scalar	0.067	-0.012	0.984	+0.003	0.984	+0.006	Yes
Strict	0.068	+0.001	0.983	-0.001	0.984	0	Yes
Federal Poverty Level							
Configural	0.085	--	0.979	--	0.974	--	
Metric	0.081	-0.004	0.978	-0.001	0.976	+0.002	Yes
Scalar	0.068	-0.013	0.981	+0.003	0.983	+0.007	Yes
Strict	0.080	+0.012	0.973	-0.008	0.977	-0.006	No
Maternal Education							
Configural	0.084	--	0.980	--	0.975	--	
Metric	0.079	-0.005	0.980	0	0.978	+0.003	Yes
Scalar	0.065	-0.014	0.983	+0.003	0.985	+0.007	Yes
Strict	0.069	+0.004	0.980	-0.003	0.980	-0.005	Yes

Developmental Level							
Configural	0.073	--	0.984	--	0.980	--	
Metric	0.068	-0.005	0.985	+0.001	0.983	+0.003	Yes
Scalar	0.056	-0.012	0.988	+0.003	0.988	+0.005	Yes
Strict	0.044	-0.012	0.992	+0.004	0.993	-0.005	Yes

[1]Configural: model with the same number of factors and the same set of zero factor loadings in all groups
[2]Metric: model where factor loadings are held equal across groups
[3]Scalar: model where factor loadings and intercepts/thresholds are held equal across groups
[4]Strict: model where factor loadings, intercepts/thresholds and residual variance are held equal across groups
RMSEA: root-mean-squares-error of approximation; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index

4.4 Discussion

Overall, the OARS 12-item demonstrated acceptable levels of construct validity across multiple methods of analysis. The correlation matrix results indicated that there is strong correlation between the recognized key ASD domains of social interaction impairment, social communication impairment, and restricted and repetitive behaviors as measured by the OARS compared to commonly used gold-standard ADOS and ADI-R assessments. This provides evidence that what is measured as "ASD characteristics" by the OARS aligns well with what is measured by other well-validated instruments. One explanation for the weaker correlation between OARS and ADI-R domains, compared with OARS and ADOS, is the mismatch between an observation-based OARS and an interview-based ADI-R. This administration method difference also may explain the weaker correlation shown between the ADOS and ADI-R, both considered highly valid tools; although the correlations shown in SEED are lower than reported in other studies (Risi et al. 2006). An additional explanation for the stronger correlation between OARS and ADOS as opposed to ADI-R may be that clinicians who administered the OARS also administered the ADOS and thus included some ADOS-based observations in OARS judgements, which is the primary limitation of this study.

The results of the oblique rotated EFA support a 2-factor structure, in line with the DSM-5 combination of social interaction and communication into a single domain rather than the DSM-4 separation of social interaction and communication. The low loading of OARS item 1 suggests that it is not as meaningful a measure of ASD constructs as other OARS items and should be examined for improvement. The strict measurement invariance found for all socio-demographic and developmental subgroups except poverty suggests that the OARS items are measuring underlying ASD constructs

in similar ways across different groups of children. Despite not meeting criteria for the strictest measurement invariance, OARS items did meet requirements for scalar invariance among the FPL subgroup, which still indicates similar factor loadings and intercepts across different income groups. The strict invariance across the developmental level subgroups suggests that the lower sensitivity and specificity found in the previous chapter's analysis was not the result of differential measurement of underlying ASD constructs between developmental groups. Prior analyses of measurement invariance in other ASD measurement tools such as the parent-report Social Responsiveness Scale also found a range from metric to strict invariance across demographic subgroups, indicating minimal difference in measurement across groups (Frazier et al. 2014). Measurement invariance across income subgroups has not been reported for the ASD-specific tools we examined. Very few studies have examined the construct validity of an ASD tool across socio-demographic subgroups, an oversight which limits the depth of research available on measurement tool performance (McConachie et al. 2015).

As mentioned previously, the primary limitation of this work within the SEED sample is the prior administration of ADOS and ADI-R by the same clinicians administering the OARS. This clearly may have led to stronger correlations between OARS and these tools than would exist otherwise. Yet, because the OARS has rarely been administrated in the same set of children with ADOS and ADI-R information, and such comparisons have not been reported in the literature, this is an important first step towards validating the OARS. Future work would need to administer the OARS specifically as a screener prior to ADOS and ADI-R to evaluate its properties more fairly.

Additionally, the OARS 12 item tool was only administered to participants undergoing ASD evaluation having initially screened positive for ASD risk, to the exclusion of the majority of the general population-based controls. Finally, the underrepresentation of below average developmental level and non-white racial/ethnic groups limits the generalizability of these findings. The OARS should be further examined in a sample with a greater representation of children with one or more developmental delays.

Future research should examine the OARS 12-item alongside other ASD measurement tools in low-resource clinical settings to assess performance in diverse and dynamic populations, and without clinician overlap. Feasibility of OARS administration should also be examined, potentially using mixed methods approaches, to ensure that the tool is a good fit for the clinical settings in greatest need of additional no-cost instruments. Overall, these construct validity analyses build upon the previous examination of OARS performance and suggest that this measurement tool has sufficient evidence to warrant further study.

4.5 References

American Academy of Pediatrics. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics, 118*(1), 405-420.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (5th ed.). Arlington, VA: American Psychiatric Association.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, *56*(2), 81.

Centers for Disease Control and Prevention (CDC). (2015). *Autism spectrum disorder: Screening and diagnosis for health care providers.* Retrieved August 8, 2016, from http://www.cdc.gov/ncbddd/autism/hcp-screening.html

Centers for Disease Control and Prevention (CDC). (2018). *Study to explore early development (SEED).* Retrieved 8/10, 2018, from https://www.cdc.gov/ncbddd/autism/seed.html

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

de Bildt, A., Sytema, S., Zander, E., Bölte, S., Sturm, H., Yirmiya, N., et al. (2015). Autism diagnostic interview-revised (ADI-R) algorithms for toddlers and young preschoolers: Application in a non-US sample of 1,104 children. *Journal of Autism and Developmental Disorders, 45*(7), 2076-2091.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Sage publications.

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of

quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*, *18*(1), 31-44.

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics, 63*(5), 851-859.

Hagan, J. F., Shaw, J. S., & Duncan, P. M. (2007). *Bright futures: Guidelines for health supervision of infants, children, and adolescents* (4th ed.). Elk Grove Village, IL: American Academy of Pediatrics.

Jiang, G., Mai, Y., & Yuan, K. H. (2017). Advances in Measurement Invariance and Mean Comparison of Latent Variables: Equivalence Testing and A Projection-Based Approach. *Frontiers in Psychology*, *8*, 1823.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685.

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal, 14*(3), 234-252.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, *30*(3), 205-223.

McConachie H, Parr JR, Glod M, et al. (2015) Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment 19*(41).

Mullen, E. M. (1995). *Mullen Scales of Early Learning* AGS Circle Pines, MN.

OSU Research Unit on Pediatric Psychopharmacology (OSU RUPP). (2005). *OSU autism rating scale-DSM 4 (OARS-4)*. Columbus, Ohio: OSU Research Unit on Pediatric Psychopharmacology.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90.

Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. Journal of the American Academy of Child and Adolescent Psychiatry. 2006;45:1094–1103.

Rutter M, Bailey A, Lord C. Social Communication Questionnaire (SCQ). Los Angeles: Western Psychological Services; 2003.

Schendel, D. E., DiGuiseppi, C., Croen, L. A., Fallin, M. D., Reed, P. L., Schieve, L. A., et al. (2012). The study to explore early development (SEED): A multisite epidemiologic study of autism by the centers for autism and developmental disabilities research and epidemiology (CADDRE) network. *Journal of Autism and Developmental Disorders, 42*(10), 2121-2140.

Schulz E. (2016). *Executive summary: Capacity building for ASD and DD: Community Based autism liaison and training (CoBALT) project in Arkansas.*

U.S Department of Health and Human Services. (2018). *HHS 2012 poverty guidelines.* Retrieved 7/1, 2018, from https://aspe.hhs.gov/2012-hhs-poverty-guidelines

Warren, Z., Stone, W., & Humberd, Q. (2009). A training model for the diagnosis of autism in community pediatric practice. *Journal of Developmental & Behavioral Pediatrics, 30*(5), 442-446.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-11.

Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., DiGuiseppi, C., et al. (2015). Autism spectrum disorder symptoms among children enrolled in the study to explore early development (SEED). *Journal of Autism and Developmental Disorders, 45*(10), 3183-3194.

Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., et al. (2015). Using standardized diagnostic instruments to classify children with autism in the study to explore early development. *Journal of Autism and Developmental Disorders, 45*(5), 1271-1280.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., et al. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136 Suppl 1*, S41-59.

CHAPTER 5. ASSOCIATION OF SOCIO-DEMOGRAPHIC CHARACTERISTICS WITH TIMING OF AUTISM SPECTRUM DISORDER IDENTIFICATION FROM THE AUTISM AND DEVELOPMENTAL DISABILITIES MONITORING NETWORK 2006-2012

5.1 Introduction

The timely identification of Autism Spectrum Disorder (ASD) is one of the primary factors determining whether children receive access to early intervention services. When accessed within the critical developmental window between ages 2 and 5, early intervention can help reduce impairment related to ASD and promote the social development necessary for later school success (Warren et al. 2011). Beyond its importance for individual child development, early identification of ASD also benefits communities by informing clinical and educational resource allocation and planning, policymaking, and the study of ASD epidemiology. The most recent estimate of median age at ASD identification in the U.S from the 2016 Autism and Developmental Disabilities Monitoring Network (ADDM) cohort is 4.3 years (Baio et al. 2018). Prior work in the ADDM network, as well as in smaller community-based studies, suggests that age at ASD identification varies across community subgroups by factors such as sex, race, geographic location, and socioeconomic status (Shattuck et al. 2009; Mandell et al. 2002; Jo et al. 2010).

The current pathway for the identification of ASD in the U.S. is a multi-step process involving community-based screening in pediatric primary care, specialist referral, and multidisciplinary approaches for diagnostic evaluation. The American Board of Pediatrics recommends that all children be screened with ASD-specific tools at 18 and 24 months in conjunction with ongoing developmental screening and surveillance (Zwaigenbaum et al. 2015; Hagan et al. 2017). Once the threshold for concern is met on

the selected screening measure, such as the Modified Checklist for Autism in Toddlers (M-CHAT-R), community physicians can refer families to seek formal evaluation from local developmental specialists. Commonly utilized developmental specialists include developmental pediatricians, psychologists, and speech language pathologists in hospital and community clinic settings. Programs funded through state Departments of Education, such as Infants and Toddlers for children younger than age 3, also provide limited initial evaluation in addition to early intervention services. The process from initial screening to receipt of diagnosis has been described by some caregivers as a "diagnostic odyssey" and can take anywhere from an estimated two months to over a year (Lappé, et al. 2018; Gordon-Lipkin et al, 2017; Bisgaier et al, 2011). Delays in identification have been attributed to a limited supply of clinical expertise and resources, fragmented referral and care coordination systems, increasing numbers of children requiring evaluation, and healthcare access inequalities at each step on the outlined pathway (Gordon-Lipkin et al. 2017; Bisgaier et al. 2011; Fenikile et al. 2015).

Socio-demographic factors often associated with disparities in healthcare quality and access in the U.S such as geographic location, race/ethnicity, and socio-economic status may also play a role in ASD identification timing. In the ADDM network, median age at identification (AAI) varies significantly by state, ranging from 3.3 years in North Carolina to 4.9 years in Arkansas in the most current 2016 study year (SY) (Baio et al. 2018). This variation is due in part to differences in the availability of health and educational records to determination of ASD status across sites in the ADDM network, but is also reflective of state and community policies, practices, and available clinical infrastructure. Work in the ADDM network by Shattuck and colleagues (2009) using the

2002 surveillance year cohort found that Asian and non-Hispanic black children had the youngest ages of identification at 5.2 and 5.3 years, while white, non-Hispanic children had a median AAI of 5.7 years. Hispanic children had the oldest AAI at 6.7 years. A smaller study using Philadelphia Medicaid claims data found an average age of ASD diagnosis for non-Hispanic black children of 7.9 years, much later than found in the ADDM network data (Mandell et al. 2002). Analyses using the National Survey on Children with Special Healthcare Needs found that children living in households with self-reported incomes <200 % of the federal poverty level (FPL) were 44% more likely to have a later diagnosis than their high income peers (Jo et al. 2015). Within the ADDM network, later AAI can influence prevalence estimates. Prior analyses have found lower estimated prevalence of ASD by age 8 among children with lower SES (Durkin et al. 2017; Maenner et al 2009)

The purpose of this study is to examine factors related to the timing of ASD identification using ADDM network data from 2006-2012. Building on prior examination of the influence of socio-demographic variables on ASD identification timing, we used available census-linkage in the ADDM network to add community-level poverty as a variable of interest. We hypothesized that median age at ASD identification would decrease across cohort years. We further predicted that lower income communities would have a later AAI than higher income communities after controlling for other child and family-level factors. By examining how age at ASD identification has changed over time and what factors are related to timing delays in the U.S, we can better inform and strengthen the current identification infrastructure.

5.2 Methods

Surveillance and abstraction methodology:

The Autism and Developmental Disabilities Monitoring (ADDM) Network is a collaborative, active surveillance system that provides estimates of the prevalence of ASD among children age 8 years in the U.S. (Baio, 2018). Surveillance for the ADDM network is conducted in two phases: review and abstraction of child records and systematic clinical review for determination of ASD status. Information relevant to the determination of ASD case status is abstracted from data sources either classified as health, including diagnostic and developmental assessments from professionals such as psychologists, or educational, including evaluations to determine eligibility for special education services. All ADDM network sites have agreements to review and abstract diagnostically relevant information from health records, but not all have equivalent permissions for accessing educational records. School and health records are reviewed and abstracted by trained staff at each of the study sites. The abstracted data is then de-identified and reviewed by clinical experts with standardized training to determine ASD case status based on DSM diagnostic criteria. Review of data for this analysis utilized DSM-IV diagnostic criteria. Specific site-level information, abstraction and clinical review procedures and considerations are outlined in ADDM Network surveillance summaries (Baio, 2018).

In addition to diagnostic information, socio-demographic data is also collected on children included in ADDM cohorts. Birth certificate data is obtained from state vital records and includes both child and maternal information. If child race is not available in health or educational source records, it is determined based on race/ethnicity of one or both parents listed on the birth certificate. Maternal educational attainment and age at

child's birth are also collected from birth certificates. Census data has been made available for linkage to ADDM network data based on zip code given within each site.

Study Sample:

Data for this analysis come from ADDM network 2006-2012 SYs and include children age 8 years whose parents or guardians resided in Alabama, Arkansas, Arizona, Colorado, Georgia, Maryland, Missouri, North Carolina, New Jersey and Utah. Children were included in the analysis if they had a final case status of ASD and at least partial birth certificate and census level information available. Children were excluded from review if they came from sites that did not have census level linkage. The final analytic sample included 13,731 children from cohort years 2006-2012.

Variables:

The primary variable of interest in this analysis was age at first autism identification, as indicated in health or educational records by a diagnosis from a formal ASD evaluation, eligibility for ASD-related special education services, or use of ICD-9 codes for ASD. Children without a recorded ASD diagnosis through age 8, but who met ADDM criteria for ASD based on clinician review have traditionally been censored from CDC estimates of age at identification (Shattuck et al, 2009). In this analysis, these previously unidentified children were considered identified at age 8 based on ADDM clinician review. Demographic information such as child race and sex as well as maternal education and age at child's birth were collected from birth certificate data. Median household income and percentage of households with children under the age of 18 living below the federal poverty line for each were obtained from census data. The percentage of households in poverty was analyzed based on federal definition of a "poverty area",

meaning greater than or equal to 20% of households living below the poverty line (U.S. Census Bureau, 2018).

Statistical Analysis:

Two types of survival analysis were conducted to explore influences of socio-demographic variables on timing of ASD identification. Timing of ASD diagnosis was first examined by using Kaplan-Meier survival curves to estimate the cumulative probability of reaching a given age without receiving an ASD diagnosis. The median age of identification was determined from the curves, based on age at which 50% of the sample has been identified as having ASD. Median AAI was assessed for the total sample as well as across socio-demographic, developmental, site and surveillance subgroups. Kaplan-Meier derived median AAI across subgroups was calculated both inclusive and exclusive of children not-yet-identified by age 8. The equality of the survivor functions was tested within each subgroup using a Tarone-Ware chi-squared test (Shattuck et al 2009; Rosner, 2011). A Tarone-Ware chi-square was also used to test for a linear trend in median AAI across cohort years 2006-2012.

To examine how age at ASD identification is associated with one or more covariates of interest, multivariate parametric survival analyses were conducted. This accelerated failure time (AFT) approach to survival analysis produced time ratios that are more easily interpretable than a traditional hazard ratio. Time ratios can be interpreted as the extent to which a predictor speeds up or delays ASD identification, relative to the selected reference category. After checking that the data violated the assumption of proportional hazards, a series of increasingly complex multivariate models were fit and included mixed-effects models which nested child, family, surveillance characteristics

within “poverty area” and site levels. Model fit was assessed using likelihood ratio tests comparing each expanded model to the previous model. Adjusted AAI reported in months was calculated from the final model and compared to reference levels to calculate month-level differences in identification timing across subgroups.

Reference categories were selected based on proximity to the overall group median AAI. For example, as the largest race/ethnicity subgroup, white, non-Hispanic children had the closest Kaplan-Meier derived median AAI to the total sample estimate. Therefore, the white non-Hispanic group was used as the reference level for the race/ethnicity subgroup. For cohort year, 2006 was used as the reference, not because it was closest to the total median, but rather to enable examination of change in median AAI over time. Stata 15 was used for all analyses.

5.3 Results

Sample description

Table 1 shows descriptive statistics for the 13,731 children included in this analysis. The male to female ratio of the sample was 4.7:1, reflective of the higher prevalence of ASD among males. The majority of the sample (57.9%) was white, non-Hispanic. Approximately half of the sample had IQ at or above 70, with a further 21% of the sample missing information on IQ in abstracted data. The distribution of maternal age and education at birth of the ADDM identified child is shown in Table 1, with roughly a quarter of the sample missing information for both variables after abstraction. Approximately 30% of children lived in census-tracts, hereafter referred to as “communities”, considered to be poverty areas based on federal guidelines. The number of children that contributed to this analysis from each study site varied, with Alabama accounting for the lowest percentage (3.7%) of children and Georgia the highest (18.9%).

With regard to type of records abstracted for ADDM clinician review, only 37.8% of children had data from both clinical and educational records available.

There was statistically significant variation in the distribution of child and family level characteristics between children with and without documented ASD diagnosis by time of ADDM record abstraction, those who were considered "not-yet-identified" (table 1). Of the total sample, 30% were not-yet-identified and the percentage of children not-yet-identified in each SY decreased from 35% (717 of 2,033) in the 2006 SY to 25% (995 of 3,873) in 2012. These children were more likely to identify as African-American or Hispanic, have IQs greater than 70 and be born to mothers with less than a high school education. With regard to community-level poverty, children not-yet-identified by age 8 were also more likely to belong to a community with lower median income and a higher percentage of households in poverty levels than their peers. The greatest source of difference between these two groups came from the examining record type available for review.

Median age at identification

In the total sample inclusive of children not-yet-identified at age 8, the median age at ASD identification was 5.8 years (Table 1). Excluding the 30% of children not-yet-identified by age 8, this median age drops to 4.3. In both samples, males and females had equivalent ages at identification and survivor functions, while all other subgroups differed significantly across subgroup levels. Hispanic children had the latest median AAI, 6.3 and 4.6 years, compared to their non-Hispanic peers. Children with IQ below 70 had the earliest a median AAI, at least a year earlier than those without potential intellectual disability in both samples. Children born to teenage mothers had the latest

AAI, 6.4 and 4.8 years, compared to other maternal age groups. Median AAI was negatively correlated with maternal education level and community-level household income.

Age at identification decreased over time, with 2012 having the earliest ages in both samples. In the sample inclusive of children not-yet-identified at age 8, the median AAI was 5.3 years, a full year earlier than the estimated age for the 2006 cohort, while in the sample excluding these children, there was a difference of 3 months. In both samples, New Jersey had the earliest AAI at 4.8 and 3.9 years. Arkansas, 6.1 and 5.0 years and Arizona, 7.1 and 4.8 years, had the latest ages at identification. In both samples, children with both educational and clinical records available for review had earlier median identification age than those with only data from one record source. Differences in AAI across levels of each subgroup were smaller in the sample exclusive of children not-yet-identified compared to the same levels in the inclusive sample.

Table 1. Socio-demographic Variable Distribution for Children Age 8 Meeting Case Criteria in the ADDM Network from 2006-2012

Variable	Total Sample	Documented ASD Diagnosis by age 8	No Documented ASD Diagnosis by age 8	
	N (%)	N(%)	N(%)	χ^2, (df), p-value
All ASD Cases	13,731 (100)	9,587 (69.8)	4,144 (30.2)	
Child and Family Characteristics				
Sex				
Male	11,330 (82.5)	7,935 (82.7)	3,396 (81.9)	1.34 (1), 0.25
Female	2,401 (17.5)	1,652 (17.3)	748 (18.1)	
Race/Ethnicity				
White, non-Hispanic	7,952 (57.9)	5,739 (59.9)	2,212 (53.4)	82.7 (4), <0.001
African-American, Non-Hispanic	2,763 (20.1)	1,853 (19.3)	911 (22.0)	
Other race/multiracial, Non-Hispanic	1,077 (7.3)	691 (7.2)	316 (7.6)	
Hispanic, regardless of race	1,769 (12.9)	1,113 (11.6)	656 (15.8)	
Missing race/ethnicity	240 (1.8)	191 (1.9)	49 (1.2)	
Cognitive Status				
Above ID Range IQ >70	7,283 (53.0)	4,747 (49.5)	2,535 (61.2)	172.4 (2), <0.001
At or Below ID Range IQ ≤70	3,563 (26.0)	2, 605 (27.2)	959 (23.1)	
IQ information missing	2,885 (21.0)	2,23 (23.3)	650 (15.7)	
Maternal Age at Birth				
<20 years	630 (4.6)	403 (4.2)	227 (5.5)	18.3 (4), <0.001
20-29 years	4,639 (33.8)	3,210 (33.5)	1,429 (34.5)	
30-39 years	4,561 (33.2)	3,265 (34.1)	1,396 (31.3)	
40 years and greater	417 (3.0)	286 (2.9)	131 (3.2)	
Maternal age missing	3,484 (25.4)	2,423 (25.3)	1,061 (25.6)	

Maternal Education at Birth				
<12 years	1,242 (9.0)	762 (7.9)	480 (11.6)	65.0 (5), <0.001
12 years	2,689 (19.6)	1,832 (19.1)	857 (20.7)	
1-3 years college	2,499 (17.9)	1,763 (18.4)	696 (16.8)	
4 years college	2,282 (16.6)	1,659 (17.3)	623 (15.0)	
5 years college and greater	1,388 (10.1)	1,011 (10.6)	377 (9.10)	
Education missing	3,671 (26.7)	2,560 (26.7)	1,111 (26.8)	
Census Tract Variables				
Community Median Income				43.2 (5), <0.001
<20,000	103 (0.8)	68 (0.7)	35 (0.8)	
20,000-39,000	2,139 (15.6)	1,413 (14.7)	726 (17.4)	
40,000-59,000	4,194 (30.5)	2,864 (29.8)	1,330 (32.1)	
60,000-79,000	3,298 (24.0)	2,320 (24.2)	979 (23.6)	
80,000-99,000	2,197 (16.0)	1,585 (16.5)	613 (14.8)	
100,000 and greater	1,798 (13.1)	1,337 (13.9)	461 (11.1)	
Community Percentage of Households in Poverty				
<10%	6,557 (47.8)	4,703 (49.1)	1,852 (44.6)	
10-19%	3,092 (22.5)	2,166 (22.6)	927 (22.3)	
[a]20-29%	1,843 (13.4)	1,256 (13.1)	588 (14.2)	
30-39%	1,075 (7.8)	697 (7.3)	378 (9.1)	
40-49%	656 (4.8)	425 (4.4)	231 (5.6)	
50% and greater	508 (3.7)	340 (3.5)	168 (4.1)	
Surveillance Variables				
Cohort Year				
2006	2,033 (14.8)	1,316 (13.7)	717 (17.3)	
2008	2,824 (20.6)	1,864 (19.4)	960 (23.2)	
2010	5,001 (36.4)	3,529 (36.8)	1,472 (35.5)	
2012	3,873 (28.2)	2,878 (30.0)	995 (24.0)	

Sites				433.5 (9), <0.001
Alabama	506 (3.7)	337 (3.5)	169 (4.1)	
Arkansas	782 (5.7)	551 (5.8)	231 (5.6)	
Arizona	2,035 (14.8)	1,208 (12.6)	827 (19.9)	
Colorado	518 (3.8)	317 (3.3)	201 (4.9)	
Georgia	2,598 (18.9)	1,718 (17.9)	880 (21.2)	
Maryland	1,359 (9.9)	1,079 (11.3)	280 (6.8)	
Missouri	1,334 (9.7)	1,085 (11.3)	249 (6.0)	
North Carolina	2,062 (15.0)	1,296 (13.5)	766 (18.5)	
New Jersey	1,636 (11.9)	1,262 (13.2)	374 (9.0)	
Utah	901 (6.6)	734 (7.7)	167 (4.0)	
Record Type Reviewed				535.3 (2), <0.001
Health Records only	4,197 (30.4)	2,844 (29.7)	1,353 (32.6)	
Educational Records only	4,363 (31.8)	2,575 (26.8)	1,788 (43.2)	
Health and Educational Records	5,171 (37.8)	4,168 (43.5)	1,003 (24.2)	

a. "poverty area" considered >20% of residents living below FPL

Table 2. Kaplan-Meier Estimates of Median Age at ASD Identification

Variable	Median Age at Identification Including Children Not-Yet-Identified by Age 8		Median Age Identification Excluding Children Not-Yet-Identified by Age 8	
	Years (95% CI)	**χ^2, (df), p-value**	**Years (95% CI)**	**χ^2, (df), p-value**
All ASD Cases	5.8 (5.7, 5.9)	-	4.3 (4.3, 4.4)	-
Child and Family Characteristics				
Sex		0.22 (1), 0.63		1.45(1), 0.23
Male	5.8 (5.5, 6.0)		4.3 (4.2, 4.4)	
Female	5.8 (5.7, 5.9)		4.4 (4.3, 4.5)	
Race/Ethnicity		60.2 (4), <0.001		10.2 (4), 0.04
White, non-Hispanic	5.7 (5.6, 5.8)		4.3 (4.3, 4.4)	
African-American, Non-Hispanic	5.9 (5.7, 6.2)		4.3 (4.2, 4.5)	
Other race/multiracial, Non-Hispanic	5.6 (5.3, 5.9)		4.3 (4.2, 3.6)	
Hispanic, regardless of race	6.3 (5.9, 6.5)		4.6 (4.4, 4.8)	
Missing race/ethnicity	5.6 (5.1, 6.4)		4.9 (4.5, 5.2)	
Cognitive Status		402.5 (2), <0.001		343.2 (2), <0.001
Above ID Range IQ >70	6.6 (6.5, 6.8)		4.9 (4.8, 5.0)	
At or Below ID Range IQ ≤70	4.9 (4.8, 5.0)		3.8 (3.7, 3.9)	
IQ information missing	5.0 (4.8, 5.2)		4.0 (3.8, 4.1)	
Maternal Age at Birth		54.3 (4), <0.001		73.6 (4), <0.001
<20 years	6.4 (6.0, 6.9)		4.8 (4.4, 5.0)	
20-29 years	5.9 (5.8, 6.0)		4.4 (4.3, 4.6)	
30-39 years	5.3 (5.2, 5.5)		4.0 (3.9, 4.6)	
40 years and greater	5.7 (5.3, 6.5)		4.0 (3.6, 4.6)	
Maternal age missing	6.1 (5.9, 6.3)		4.7 (4.6, 4.8)	
Maternal Education at Birth		111.4(5), <0.001		80.7 (5), <0.001
<12 years	6.7 (6.3, 7.1)		4.8 (4.5, 4.8)	
12 years	5.9 (5.7, 6.2)		4.5 (4.5, 4.8)	
1-3 years college	5.7 (5.4, 5.9)		4.3 (4.1, 4.4)	
4 years college	5.3 (5.2, 5.5)		4.0 (3.8, 4.2)	
5 years college and greater	5.2 (5.0, 6.2)		3.9 (4.5, 4.75)	
Education missing	6.0 (5.9, 6.2)		4.7 (4.5, 4.4)	
Census Tract Variables				
Community Median Income		78.17 (5), <0.001		35.8 (5), <0.001
<20,000	6.5 (5.5, 7.4)		5.1 (4.5, 5.5)	
20,000-39,000	6.3 (6.0, 6.4)		4.6 (4.4, 4.8)	
40,000-59,000	5.9 (5.8, 6.1)		4.5 (4.4, 4.6)	
60,000-79,000	5.8 (5.6, 6.0)		4.3 (4.2, 4.4)	
80,000-99,000	5.8 (5.5, 5.9)		4.3 (3.8, 4.1)	
100,000 and greater	5.1 (5.0, 5.4)		4.0 (3.8, 4.1)	
Community Percentage of Households in Poverty		56.8 (5), <0.001		19.5 (5), 0.002
<10%	5.6 (5.4, 5.7)		4.2 (4.1, 4.3)	
10-19%	5.8 (5.7, 5.9)		4.4 (4.3, 4.6)	

[a]20-29%	5.9 (5.7, 6.25)		4.6 (4.4, 4.8)	
30-39%	6.3 (6.0, 6.8)		4.6 (4.4, 4.8)	
40-49%	6.2 (5.8, 6.6)		4.6 (4.4, 4.9)	
50% and greater	6.2 (5.6, 6.7)		4.4 (4.1, 4.8)	
Surveillance Variables				
Cohort Year		105 (3), <0.001		21.8 (3), <0.001
2006	6.3 (6.2, 6.6)		4.5 (4.3, 4.7)	
2008	6.3 (6.1, 6.6)		4.6 (4.4, 4.8)	
2010	5.8 (5.6, 5.8)		4.4 (4.3, 4.5)	
2012	5.3 (5.1, 5.4)		4.2 (4.1, 4.3)	
Sites		405(9), <0.001		110.2 (9), <0.001
Alabama	6.1 (5.8, 6.7)		4.5 (4.2, 4.9)	
Arkansas	6.1 (5.8, 6.4)		5.0 (4.8, 5.2)	
Arizona	7.1 (6.8, 7.4)		4.8 (4.7, 4.9)	
Colorado	6.9 (6.4, 7.5)		4.9 (4.6, 5.2)	
Georgia	6.0 (5.8, 6.2)		4.3 (4.2, 4.4)	
Maryland	5.7 (5.4, 5.8)		4.7 (4.4, 5.0)	
Missouri	5.1 (4.9, 5.3)		4.3 (3.9, 4.5)	
North Carolina	6.3 (6.0, 6.7)		3.9 (3.8, 4.1)	
New Jersey	4.8 (4.7, 5.0)		3.9 (3.8, 4.1)	
Utah	4.9 (4.7, 5.2)		4.3 (4.2, 4.5)	
Record Type Reviewed		670 (2), <0.001		131 (2), <0.001
Health Records only	6.1 (5.9, 6.3)		4.3 (4.2, 4.5)	
Educational Records only	7.1 (7.0, 7.3)		4.1 (4.0, 4.2)	
Health and Educational Records	4.8 (4.7, 4.9)		4.1 (4.0, 4.2)	

Multivariate survival analysis

Controlling for other child and family level covariates, later AAI was significantly associated with African-American or Hispanic race/ethnicity and maternal education level of high school or less (Table 3). Earlier AAI was significantly associated with IQ below 70 or missing information on IQ, maternal age at birth 30-39, and college or greater maternal education. Model 2 adds study year as a covariate, and the same significant associations between child and family level covariates and timing are found as in model 1. Belonging to the 2010 or 2012 cohort was significantly associated with earlier AAI.

Table 3. Multivariate Accelerated Failure Time Survival Models of Age at ASD Identification for Children Age 8 Meeting Case Criteria in the ADDM Network from 2006-2012

	Time Ratios			
	Model 1	**Model 2**	**Model 3**	**Model 4**
Fixed-Effects Model				
Child and Family Characteristics				
Sex				
Male	0.98	0.98	0.98	0.98
Female	--	--	--	--
Race/Ethnicity				
White, non-Hispanic	--	--	--	--
African-American, Non-Hispanic	1.07***	1.07***	1.07***	1.05***
Other race/multiracial, Non-Hispanic	1.03	1.03*	1.03	1.02
Hispanic, regardless of race	1.07***	1.08***	1.07***	1.06***
Missing race/ethnicity	1.01	1.01	1.01	1.05
Cognitive Status				
Above ID Range IQ >70	--	--	--	
At or Below ID Range IQ ≤70	0.78***	0.78***	0.81***	0.80***
IQ information missing	0.80***	0.80***	0.77***	0.80***
Maternal Age at Birth				
<20 years	1.02	1.02	1.02	1.01
20-29 years	--	--	--	--
30-39 years	0.96**	0.96***	0.97**	0.97*
40 years and greater	1.01	1.02	1.02	1.02
Maternal age missing	1.04	1.04	1.02	1.02
Maternal Education at Birth				
<12 years	1.12***	1.12***	1.12***	1.11***
12 years	1.04**	1.05***	1.05**	1.04**
1-3 years college	--	--	--	--
4 years college	0.95**	0.95***	0.96*	0.97*
5 years college and greater	0.94**	0.95***	0.96*	0.96*
Education missing	0.99	0.98	0.99	0.99
Surveillance Variables				
Cohort Year				
2006		--	--	--
2008		0.98	0.99	0.99
2010		0.92***	0.95***	0.95***
2012		0.88***	0.90***	0.91***
Record Type				
Health Records only			1.22***	1.23***
Educational Records only			1.31***	1.31***
Both Health and Educational Records			--	--
Constant	72.0	77.4	65.4	65.9
Random Effects				
Variance attributed to poverty area				~0
Variance attributed to study site				0.005
Log Likelihood	-67093.1	-67031.1	-66685.5	-66614.6
Likelihood Ratio Test p value	--	<0.001	<0.001	<0.001

*P=0.05, **P=0.01, ***P=0.001

Model 3 added record source to child, family, and cohort variables. In this complex model, the significant associations between earlier and later ages at identification with race, maternal education, IQ, and cohort year were observed. However, once adjusting for record source, cohort year 2010 also had a significantly later AAI compared to 2006. Having only health or educational records was significantly associated with later ages at identification, with 23% and 31% respective increases in age identification compared to children with both sources available. Model 4 nested the previous model within two levels of geographically defined variables: a binary "poverty area" variable indicating whether children were identified from a community with greater than 20% of households below the FPL and a study site variable to indicate in which ADDM network site children resided. In model 4, there was little variance in these time ratios across study site or poverty area levels, with site responsible for a greater amount of variance than poverty area level.

The adjusted time ratios for model 4 were used to calculate adjusted AAI across the subgroups examined. Table 4 places these time ratios in the context of differences in AAI by months compared to reference levels of each variable. Controlling for other variables, being African-American, Hispanic or missing information on race/ethnicity was associated with a 3-month later identification compared to white, non-Hispanic children. The 31% delay in ASD identification for children with IQ below 70 in model 4 results in a difference of 13 months. Adjusting for other variables, maternal education less than high school was associated with an ASD identification 7 months later than the "some college" reference group. Membership in the 2012 cohort was associated with a 5 month earlier AAI compared to the 2006 cohort. A trend test on the adjusted median AAI

across cohort years found a significant trend towards earlier ages in later cohort years (Figure 1). Surveillance record source had the largest differences in identification age. Having only health or educational records available for review was associated with 15 and 20 month later ages respectively compared to children with both records types available for ADDM clinician review.

Table 4. **Differences in Adjusted Median Age at Identification by Months**

	Adjusted Median Age at Identification (months)	Difference from Reference Group (months)
Child and Family Characteristics		
Sex		
Male	63	-1
Female	64	--
Race/Ethnicity		
White, non-Hispanic	64	--
African-American, Non-Hispanic	67	+3
Other race/multiracial, Non-Hispanic	65	+1
Hispanic, regardless of race	67	+3
Missing race/ethnicity	67	+3
Cognitive Status		
Above ID Range IQ >70	64	--
At or Below ID Range IQ ≤70	51	-13
IQ information missing	52	-12
Maternal Age at Birth		
<20 years	64	0
20-29 years	64	--
30-39 years	62	-2
40 years and greater	66	+2
Maternal age missing	64	0
Maternal Education at Birth		
<12 years	71	+7
12 years	67	+3
1-3 years college	64	--
4 years college	62	-2
5 years college and greater	61	-3
Education missing	63	-1
Surveillance Variables		
Cohort Year		
2006	64	--
2008	63	-1
2010	60	-4
2012	59	-5
Record Type Reviewed		
Health Records only	79	+15
Educational Records only	84	+20
Both Health and Educational Records	64	--

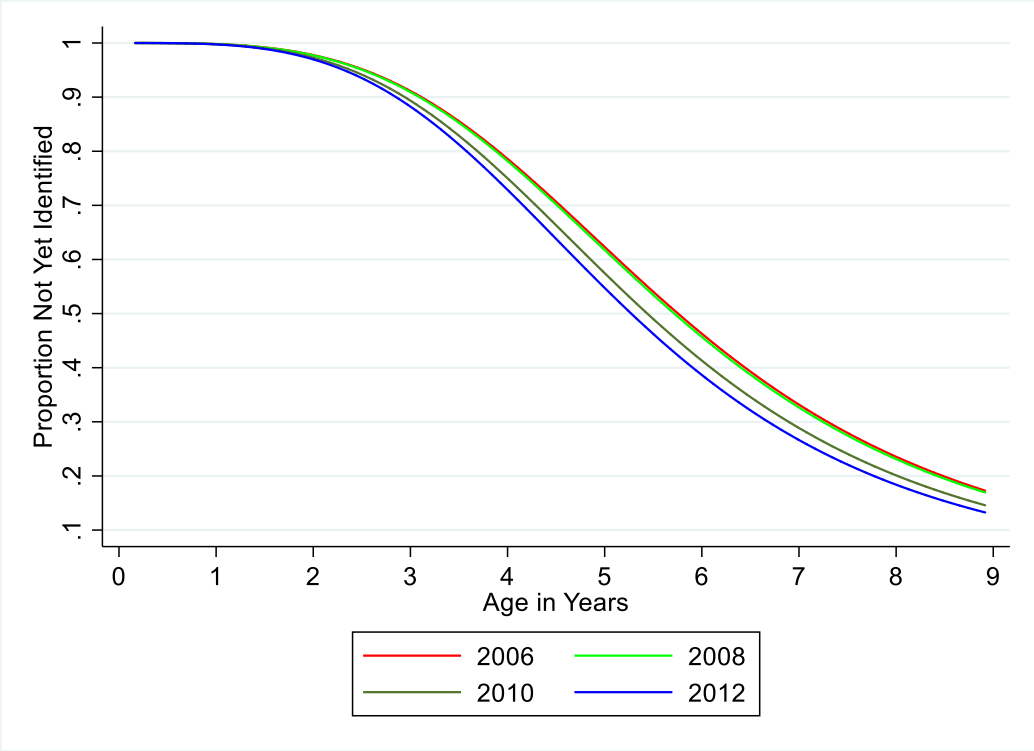


Figure 1. *Adjusted Age at ASD Identification by ADDM study year. Test for trend of survival functions across cohorts 2006-2012:* χ^2*=21.8 p=<0.001.*

5.4 Discussion

Median ages at identification reported here are later than those previously reported in CDC estimates, due to the inclusion of children who were not identified as having ASD prior to record abstraction and review by ADDM staff. A third of the total sample was not-yet-identified as ASD by age 8. Children in this group were more likely to identify as African-American or Hispanic, have IQs greater than 70, be born to mothers with less than a high school education and live in communities with greater economic disadvantage. Exclusion of children who were not-yet-identified at age 8 may artificially lower estimates of identification age in the ADDM network by failing to reflect the experience of children who have been otherwise undetected or misclassified by current clinical and educational pathways. The availability of both health and educational

records for review by ADDM staff, rather than just one source, allows for more comprehensive and complete data in obtaining these estimates.

Overall, several socio-demographic factors were shown to be significantly associated with ASD identification timing. Factors significantly associated with earlier ages at detection included lower child IQ and higher levels of maternal education. Children with IQ < 70 may have experienced delays meeting recommended early developmental milestones, resulting in earlier referrals to developmental specialists for evaluation. Higher levels of maternal education may relate to improved health literacy, agency to seek specialist referral, and awareness of milestones for motor, verbal, and social development.

Child and family level factors significantly related to later ages at detection included being of African-American or Hispanic race/ethnicity and maternal education of high school or less. Previous studies have identified lower parental awareness of ASD early signs and symptoms as well as barriers to accessing timely and appropriate pediatric care for both of these populations (Flores & Committee on Pediatric Research, 2010). While several socio-demographic factors were significantly associated with earlier or later ages at identification, the translation of these time ratios into month differences in table 4 shows that this statistical significance may not uniformly translate into clinically meaningful difference.

Our hypothesis that median age at ASD identification would decrease from 2006 to 2012 was supported by these results. Age at identification was shown to decrease from 2006 to 2012, with 2010 and 2012 cohort membership associated with statistically significant earlier ages than 2006. Our hypothesis that children in areas of higher

economic disadvantage would have a later AAI than those from higher income communities after consideration of child and family level characteristics was not supported by these results. Our findings showed approximately zero variance due to clustering on poverty and a small amount due to study site. Evidence in other areas of pediatric medicine such as vaccinations and asthma control suggests that inequities exist in healthcare access based on income (CDC, 2011; Grant et al. 2014 ;Fung et al. 2011; Larson et al. 2016). It is important to place these cohort years in the context of the ages at which children were commonly identified. Considering a median AAI for the total sample of 5.8 years, in using data from children who were age 8 at years 2006-2012, we are examining identification processes that occurred from 2001-2008. While the use of general developmental screening tools such as the Denver Developmental Screening Test first introduced in 1961 have been historically recommended for use in pediatric primary care, the American Academy of Pediatrics did not recommend the universal use of ASD-specific screening tools such as the M-CHAT at well-child visits until 2007 (Johnson, Meyers, and AAP Council on Children with Disabilities, 2007). One potential explanation for why community-level poverty did not contribute to variation in the estimation of time ratios may be that all children at relevant ages during the time period examined were simply not receiving routine ASD-specific screening and identification, regardless of the quality of healthcare accessed. These research questions should be re-examined in later ADDM cohorts to capture associations with ASD identification timing after the 2007 move for universal early ASD screening.

This study has several limitations. Primarily, the range and quality of the variables examined was limited by the availability of records from the ADDM network. Like other

forms of surveillance data, information from the ADDM network was obtained from documentation in records rather than from direct interaction or observation with children. For over have of the children studied, the availability of only one record source rather than both health and educational records reduced the completeness of the information available for ADDM abstraction and review. Additionally, ADDM procedures were not consistent across study sites or cohort years. Study sites were not always active in each cohort year, with Arizona, New Jersey and Utah joining in 2008 and Alabama not contributing to the 2012 cohort year. After recognizing the underrepresentation of racial/ethnic minorities in ADDM sites, there were greater procedural efforts to increase diversity of children included in the network occurring in later cohort years.

Median age at ASD identification timing in the U.S. has decreased over time as caregiver and clinician awareness has improved alongside the development of policies and protocols for early identification across the country. ASD identification is associated with many of the same socio-demographic variables seen to influence other well-documented health disparities in the U.S. Further examination of the role these factors play in the ASD identification timing will help to inform and improve current identification infrastructure.

5.5 References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (5th ed.). Arlington, VA: American Psychiatric Association.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C.), 67*(6), 1-23.

Bisgaier, J., Levinson, D., Cutts, D. B., & Rhodes, K. V. (2011). Access to autism evaluation appointments with developmental-behavioral and neurodevelopmental subspecialists. *Archives of Pediatrics & Adolescent Medicine, 165*(7), 673-674.

Durkin, M. S., Maenner, M. J., Baio, J., Christensen, D., Daniels, J., Fitzgerald, R., ... & Wingate, M. S. (2017). Autism spectrum disorder among US children (2002–2010): socioeconomic, racial, and ethnic disparities. *American Journal of Public Health*, *107*(11), 1818-1826.

Fenikilé, T. S., Ellerbeck, K., Filippi, M. K., & Daley, C. M. (2015). Barriers to autism screening in family medicine practice: a qualitative study. *Primary Health Care Research & Development*, *16*(4), 356-366.

Flores, G., & Committee On Pediatric Research (2010). Technical report--racial and ethnic disparities in the health and health care of children. *Pediatrics*, *125*(4), e979.

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics, 63*(5), 851-859.

Hagan, J. F., Shaw, J. S., & Duncan, P. M. (2007). *Bright futures: Guidelines for Health Supervision of Infants, Children, and adolescents* (4th ed.). Elk Grove Village, IL.: American Academy of Pediatrics.

Jo, H., Schieve, L. A., Rice, C. E., Yeargin-Allsopp, M., Tian, L. H., Blumberg, S. J., ... & Boyle, C. A. (2015). Age at autism spectrum disorder (ASD) diagnosis by race, ethnicity, and primary household language among children with special health care needs, United States, 2009–2010. *Maternal and Child Health Journal*, *19*(8), 1687-1697.

Johnson, C. P., Myers, S. M. & Council on Children with Disabilities (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1183-1215.

Lappé, M., Lau, L., Dudovitz, R. N., Nelson, B. B., Karp, E. A., & Kuo, A. A. (2018). The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics*, *141*(Supplement 4), S272-S279.

Mandell, D. S., Listerud, J., Levy, S. E., & Pinto-Martin, J. A. (2002). Race differences in the at diagnosis among Medicaid-eligible children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(12), 1447-1453.

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., ... & Shattuck, P. T. (2009). Racial/ethnic disparities in the identification of children with autismspectrum disorders. *American Journal of Public Health*, *99*(3), 493-498.

Maenner, M. J., Arneson, C. L., & Durkin, M. S. (2009). Socioeconomic disparity in the prevalence of autism spectrum disorder in Wisconsin. *Wisconsin Medical Journal*, *108*(5), 37-39.

Rosner, B. (2000). *Fundamentals of Biostatistics*. Duxbury, MA: Nelson Education.

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., & Baio, J. (2009). Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 474-483.

U.S Department of Health and Human Services. (2018). *HHS 2012 poverty guidelines.* Retrieved 7/1, 2018, from https://aspe.hhs.gov/2012-hhs-poverty-guidelines

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-11.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., et al. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136 Suppl 1*, S41-59.

CHAPTER 6: DISCUSSION OF RESULTS AND PUBLIC HEALTH IMPLICATIONS

6.1 Brief Summary

This dissertation seeks to advance the understanding of how socio-demographic factors are associated with ASD measurement and the timing of identification in the U.S. As discussed in Chapters 1 and 2, the current pathway for the identification of ASD in the U.S. is a multi-step process involving community-based screening in pediatric primary care, specialist referral, and multidisciplinary approaches for diagnostic evaluation. Delays and disparities in the timely and appropriate identification of ASD can be attributed in part to measurement and structural barriers along the identification pathway including: variation in measurement tool performance across demographic and developmental subgroups, high cost of per-use and other licensed tools, and a limited supply of clinical expertise and resources (Durkin et al. 2015; Rosenberg et al. 2018; Khowaja, Hazzard & Robins, 2015; Gordon-Lipkin et al. 2017). Particularly in low resource communities, these factors may prevent or delay the standardized and multidisciplinary evaluation that is needed to access early intervention services. Limited research is available on how socio-demographic factors are associated with ASD measurement and identification timing in the U.S. Bridging this gap in knowledge can help to inform solutions that increase access to screening and care and reduce disparities in age at identification.

Measurement variation across socio-demographic characteristics was examined using data from the Study to Explore Early Development (SEED). One potential barrier to identification is the cost of screening and assessment, which rely on expensive licensed

assessment tools. To potentially overcome that cost, in Chapter 3 the performance of a no-cost ASD measurement tool, the Ohio State Autism Rating Scale (OARS 12-item), was compared to current gold-standard ASD assessments in SEED, where both types of assessment were used. Cutoffs for OARS Total Symptom Count (TSC) and Total Impairment Mean (TIM) scores were determined through ROC and AUC analyses and compared across socio-demographic and developmental subgroups. Hypothesis 1, that the OARS 12-item will have sensitivity and specificity above 0.70 across all socio-demographic strata, was supported for all subgroups except for children with below average MSEL-derived development.

Chapter 4 expanded on the understanding of how the OARS 12-item measures ASD across socio-demographic and developmental subgroups by assessing the construct validity of the tool with factor analysis and measurement equivalence testing methods. Hypothesis 2a, that OARS items will cluster on a 2-factor structure, was supported by the results of exploratory factor analysis. Hypothesis 2b, that OARS items will have strict measurement invariance across socio-demographic subgroups, was supported for all socio-demographic variables except poverty. The performance of the OARS 12-item and implications for use are further discussed in more detail in the next sections (6.2 and 6.3).

The timely identification of Autism Spectrum Disorder (ASD) is one of the primary factors determining whether children receive access to early intervention services. In Chapter 5, the association of socio-demographic factors with ASD identification timing was examined using data from the Autism and Developmental Disabilities Monitoring (ADDM) network. Median age at identification for children from the 2006, 2008, 2010 and 2012 ADDM cohorts was analyzed using available census-

linkage and birth certificate date to explore the association of timing with community level poverty and other socio-demographic factors. Hypothesis 3a, that median age of identification will have significantly decreased from 2006-2012, was supported by the results of both unadjusted and adjusted survival analysis. Hypothesis 3b, that controlling for child, family, and surveillance level factors, median age at identification will vary between high and low poverty communities, was not supported by nested multi-variate survival analysis. Further discussion of the results and implications of these ADDM analyses are provided in section 6.4.

6.2 OARS Performance

The first two studies in this dissertation (presented in Chapters 3 and 4) examined the performance of the Ohio Autism Rating Scale (OARS 12-item) across socio-demographic and developmental subgroups, to set the stage for the utility of the no-cost OARS tool as a level-two screener in low-resourced populations where expensive evaluations are not always feasible. This could decrease the time to identification for children with ASD improving their chances for meaningful early intervention. Comparison of OARS "Total Symptom Count" (TSC) and "Total Impairment Mean" (TIM) scores to SEED final case status as derived by use of gold standard ADOS and ADI-R resulted in acceptably high levels of sensitivity and specificity across all socio-demographic subgroups. Generally, the TSC had higher specificity across socio-demographic subgroups, while TIM had higher sensitivity. OARS did not show a decrease in diagnostic accuracy for lower SES or racial/ethnic minority groups as has been reported by studies of similar measures (Moody et al, 2017; Scarpa et al. 2013).

Within the MSEL-derived "below average" developmental subgroup, both TSC and TIM had imprecise sensitivities and specificities with wide confidence intervals which dipped below the set 70% threshold for acceptability. At lower cutoff scores (see Chapter 3 table 5), TSC and TIM sensitivity and specificity point estimates improved but still had wide confidence intervals containing estimates below the acceptable threshold. The decreased sensitivity and specificity of the OARS for children with below average MSEL-measured development has also been found in studies examining the performance of other measures such as the Social Responsiveness Scale and the Social Communication Questionnaire among children with below average developmental level (Moody et al. 2017; Oosterling et al. 2010).

With regard to construct validity, the correlation matrix comparing OARS, ADOS and ADI-R indicated that there is strong correlation between the recognized key ASD domains of social interaction impairment, social communication impairment, and restricted and repetitive behaviors as measured by the OARS compared to commonly used gold-standard ADOS and ADI-R assessments. This provides evidence that what is measured as "ASD characteristics" by the OARS aligns well with what is measured by other well-validated instruments. This result, alongside those from Chapter 3, suggest that the OARS 12-item has psychometric properties similar to other recognized, but costly, methods of ASD measurement that allow for the identification of ASD among high risk children age 2-5.

Chapter 4 also examined how OARS items measured underlying ASD constructs through use of factor analysis. A two-factor structure was found to best fit the data, which aligns with a DSM-5 based understanding ASD. Prior to DSM-5, ASD symptoms were

divided up in DSM-4 by social reciprocity, communicative intent, and restricted and repetitive behavior (Hyman, 2013; APA 2013). DSM-5 combined the social reciprocity and communicative intent domains into a single social interaction/communication domain to However, OARS item 1 "impairment in use of nonverbal behaviors to regulate social interaction" did not load strongly on either factor, which suggests that it may not be as meaningful a measure of ASD constructs as other OARS items and should be examined for improvement. The results of the factor analysis support this 2-domain ASD characterization.

Strict measurement invariance was observed across all socio-demographic and developmental subgroups except poverty, suggesting that the OARS items are measuring underlying ASD constructs in similar ways across different groups of children. This result, considered alongside the acceptably high sensitivity and specificity of the OARS across groups, suggests OARS may be a useful tool for most children. Other studies that examined measurement invariance in tools such as the parent-report Social Responsiveness Scale also found a range from metric to strict invariance across demographic subgroups (Frazier et al. 2014). However, comparable results of measurement invariance across income subgroups for an ASD-specific tool could not be found in the literature to date. Very few studies have examined the construct validity of an ASD tool across socio-demographic subgroups, an oversight which limits the depth of research available on measurement tool performance (McConachie et al. 2015). Future studies seeking to validate ASD measurement tools should consider including a range of socio-demographic factors such as income, race/ethnicity, and developmental level so that the performance of these tools across subgroups can be assessed. The sample

populations used to develop and validate ASD measurement tool performance should also be as diverse as possible to help reduce the potential for bias in the measurement and capture any differences in ASD presentation.

6.3 OARS Considerations for Low-Resource Settings

As an observation-based, open-source tool, the OARS potentially offers a complimentary measurement approach to reduce delays in existing identification pathways. In lower resource settings where long wait times and a high volume of children requiring evaluation result in delays in accessing early intervention services, an adaptation of the OARS could be implemented to provide an intermediate measure of ASD between initial pediatric screening and formal evaluation by a team of specialists. In its current format, the OARS is written in very technical language and relies on a broad existing clinical knowledge of child development and ASD presentation. There is no guidance for clinicians on how to motivate a child to perform tasks that gauge the behaviors measured by the OARS. For example, clinicians are asked to identify whether children display "impairment in social or emotional reciprocity" in item 4 and then rate the perceived severity of the impairment with no explicit instructions on how to perform clinical interactions to observe reciprocity. In terms of the immediate usefulness of the OARS as a screening tool, this lack of instruction and reliance on prior clinical knowledge limits the feasibility of OARS in its current format. However, this same lack of formal instructions or administration method combined with the open-source access encouraged by the OARS authors also allows for a greater potential of adaptation across cultures and contexts. Currently, the cultural adaptation of existing per-use license tools

requires permission from the original authors, publisher, and associated fees. As more funding has been provided to examine ASD in traditionally under-researched communities, researchers have called for the development of open-source tools that are suitable for adaptations to best reflect their populations of interest (Durkin et al. 2015).

Future researchers wishing to adapt the OARS for use as a screening tool should follow best practices for psychometric scale development and adaptation (DeVellis, 2016). The existing OARS provides the "item pool" and response format from which screening items can be developed. Language and assumed prior knowledge of items should be appropriate to the group that will be administering the tool. For example, if the OARS will be used by psychologists as part of Infants and Toddlers programs, the item phrasing and general knowledge about development and ASD assumed will be different than if the OARS was to be used by medical assistants in federally qualified health centers. Moreover, if the OARS is to be used by community health workers or other paraprofessionals not typically involved in direct ASD observation in low resource settings, the phrasing and assumed knowledge would again differ. In line with recommendations for scale development and adaptation, any new screening tool developed from the OARS would need to be developed with input from members of the community as well as local experts on child development and ASD. Further pilot testing of an OARS-based screener should involve assessment of validity, reliability, and utility of the tool in a diverse sample, and when used specifically as a level-two screener. The use of mixed methods approaches would allow researchers to gather both quantitative and qualitative data on how the tool performs and can be improved.

6.4 ASD Identification Timing

Chapter 5 examined the median age at ASD identification in the ADDM network form 2006-2012. Median ages at identification found in this analysis were later than those previously reported in CDC estimates, due to the inclusion of children who were not identified as having ASD prior to record abstraction and review by ADDM staff. The exclusion of children who were not-yet-identified at age 8 in previous CDC estimates of median age at identification may artificially lower estimates by failing to reflect the experience of children who have been otherwise undetected or misclassified by current clinical and educational pathways. Age at identification was shown to decrease from 2006 to 2012, with 2010 and 2012 study year membership associated with statistically significant earlier ages than 2006. Factors significantly related to earlier ages at detection included lower child IQ and higher levels of maternal education. Factors significantly related to later ages at detection included being of African-American or Hispanic race/ethnicity, maternal education of high school or less and having only a single record source reviewed. Based on a review of published studies, this was the first time that race/ethnicity has been found significantly associated with ASD timing in data from the ADDM network and perhaps results from the large, multi-year sample. In future, researchers should more closely examine factors related to children being "not-yet-identified" by time of ADDM review and abstraction. State-level differences should also be examined in future within the context of variation in policies and practices for ASD identification and service provision. Researchers should further examine the association of socio-demographic characteristics with age at ASD identification in later ADDM study years, as they become available, to assess any changes in these associations for children

who were identified as part of universal ASD screening after the 2007 recommendation from the American Academy of Pediatrics.

6.5 Limitations and Strengths

This section will discuss limitations and strengths of this dissertation overall as well as those specific to each analysis. With regard to SEED-specific limitations, procedures for classification of ASD in the SEED population limits the extent to which formal validation of OARS can be accomplished in this sample and specific limitations to the evaluation of OARS performance are examined in chapter 3 and 4. While the potential for the development of an OARS-based screening tool was the primary motivation for this work, the OARS was not used as a screener in SEED, but rather as an additional rating scale to aid in evaluation. Clinicians completing the OARS 12-item had knowledge of child behavior and deficits due to prior administration of the ADOS and ADI-R. Despite this limitation, comparison of performance of the OARS to other validated tools in SEED allowed for an examination of OARS measurement properties by socio-demographic differences, which may enable future avenues for adaptation of this tool.

Limitations of the use of surveillance-level data from the ADDM network include limitations on the range and quality of the variables examined as they came from secondary sources, as well as missing data from birth certificates. A main limitation of any assessment of ADDM data is the availability of source records, with sites that are able to abstract and review both health and educational records potentially contributing more complete and accurate data than sites that rely on review of a single record source.

While the ADDM network is designed to sample a diverse section of the U.S population, it may not be fully representative of the U.S. as a whole.

In all analyses, small sample size among Asian, Pacific Islander, Native American, and multiracial groups required the collapsing of different racial groups into composite categories. Racial/ethnic categories like "Asian" and "Hispanic" represent a large variation in experience, culture, and language that were unable to be explored due to sample size constraints. Therefore, conclusions specific to these groups should be further examined in studies that are powered to assess variations in these groups.

The primary strength of this dissertation is that it contributes to the growing body of research on how ASD measurement and identification timing are associated with socio-demographic factors in the U.S. As discussed in the conceptual framework (presented in Chapter 1) and literature review (presented in Chapter 2), child, family, community and societal level factors interact to create the environment in which children move along the ASD identification pathway, a dynamic that is considered in the analyses presented throughout this dissertation. In an effort to best reflect the available economic resources of a household and approximate financially determined healthcare access, poverty level as derived from income and household membership in SEED and the community percentage of households in poverty in the ADDM study were both chosen for these analyses over measures of income alone. The inclusion of poverty as a variable of interest in all analyses marks a strength of this work, as there is limited research on ASD measurement tool performance across SES. Another strength of this dissertation is the examination of the no-cost OARS 12-item, which may allow for future adaptation and implementation of this freely available tool to assist with ASD identification in low-

resource communities. Next, analyses of ASD identification timing included child, family, cohort and surveillance site level variables in models to examine how median age at ASD identification has changed over time rather than simply using unadjusted median ages. Finally, the ADDM analyses also included children who were not identified as having ASD prior to ADDM abstraction and review. As these children are usually excluded in CDC estimates of median age at identification, Chapter 5 provides a potentially less biased and multi-year analysis of ASD identification timing in the U.S.

6.6 Conclusions

As presented in this dissertation, socio-demographic characteristics are associated with ASD measurement and identification timing in the U.S. The results of analyses in the SEED data suggest that the OARS 12-item has psychometric properties similar to other recognized, but costly, methods of ASD measurement that allow for the identification of ASD among high risk children age 2-5. Engaging in appropriate adaptation of the OARS 12-item into a suitable screening tool and piloting it in low-resource settings may be beneficial in decreasing the time to identification and early intervention for children with ASD. Based on the findings in the ADDM network, median age at ASD identification timing in the U.S. has decreased over time as caregiver and clinician awareness has improved alongside the development of policies and protocols for early identification across the country. ASD identification is associated with many of the same socio-demographic variables seen to influence other well-documented health disparities in the U.S. Further examination of how socio-demographic factors are

related to disparities in ASD measurement and identification timing will help to inform and improve the current identification infrastructure in the U.S.

6.7 References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition* (5th ed.). Arlington, VA: American Psychiatric Association.

DeVellis, R. F. (2016). *Scale development: Theory and Applications* (Vol. 26). Thousand Oaks, CA. Sage publications.

Durkin, M. S., Maenner, M. J., Baio, J., Christensen, D., Daniels, J., Fitzgerald, R., ... & Wingate, M. S. (2017). Autism spectrum disorder among US children (2002–2010): socioeconomic, racial, and ethnic disparities. *American Journal of Public Health*, *107*(11), 1818-1826.

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*, 18(1), 31-44.

Gordon-Lipkin, E., Foster, J., & Peacock, G. (2016). Whittling down the wait time: Exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. *Pediatric Clinics, 63*(5), 851-859.

Khowaja, M. K., Hazzard, A. P., & Robins, D. L. (2015). Sociodemographic barriers to early detection of autism: screening and evaluation using the M-CHAT, M-CHAT-R, and follow-up. *Journal of Autism and Developmental Disorders*, *45*(6), 1797-1808.

McConachie H, Parr JR, Glod M, et al. (2015) Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment* 19(41).

Moody, E. J., Reyes, N., Ledbetter, C., Wiggins, L., DiGuiseppi, C., Alexander, A., ... & Rosenberg, S. A. (2017). Screening for autism with the SRS and SCQ: Variations across demographic, developmental and behavioral factors in preschool children. *Journal of Autism and Developmental Disorders*, 47(11), 3550-3561.

Oosterling, I., Rommelse, N., De Jonge, M., Van Der Gaag, R. J., Swinkels, S., Roos, S., ... & Buitelaar, J. (2010). How useful is the Social Communication Questionnaire in toddlers at risk of autism spectrum disorder?. *Journal of Child Psychology and Psychiatry*, 51(11), 1260-1268.

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research and Treatment*, *2011* Article ID 874619.

Scarpa, A., Reyes, N. M., Patriquin, M. A., Lorenzi, J., Hassenfeldt, T. A., Desai, V. J., & Kerkering, K. W. (2013). The modified checklist for autism in toddlers: Reliability in a diverse rural American sample. *Journal of Autism and Developmental Disorders*, 43(10), 2269-2279.

APPENDIX

1. *Copy of Ohio State Autism Rating Scale as used in SEED*

OSU Autism Rating Scale (OARS-4) and Clinical Global Impression (CGI) Adapted for the Study to Explore Early Development (SEED)

Description

The OSU Autism Rating Scale—DSM-IV (OARS-4) and OSU Autism CGI were developed to provide four types of summary scores: (a) A weighted score based on severity of autism or autism spectrum symptoms derived from clinical interview; (b) A symptom count, based on the same interview; (c) A global severity scale for autism, which takes autism spectrum and related symptoms (e.g., compulsions, problems transitioning, SIB) into account; and (d) A global improvement scale for autism. SEED adapted (a) and (c) to provide an individual severity rating for each autism diagnostic domain and total autism impairment.

How to Rate

The OARS-4 on the next page contains the autism signs and symptoms in the DSM—IV. These should be rated with the degree of impairment the patient/client experiences for the given symptom. Generally, the symptoms should be elicited in a child observation and semi-structured interview with the subject's primary caregiver. While scoring, try to take both frequency/duration and degree of impairment into account as well as how much the item interferes with relationships, learning, and/or activities of daily living. Thus, rituals or preoccupations that severely interfere with most attempts to transition or that are present most of the time should be scored 2 or 3. Conversely, a ritual or preoccupation that is very mild, occupies little time, and interferes only mildly with daily events may be scored as 1.

The last page contains the OSU Autism Clinical Global Impression ("OSU Autism CGI") scale, which has separate subscales for symptom severity and for global improvement. These are rated in a similar way to the NIMH CGI Severity scale, but it is focused on autism spectrum symptoms. Symptoms frequently associated with autism spectrum—such as compulsions, hyperactivity, and self injury—should be considered even if not listed in the DSM-IV symptoms. If autism symptoms are not witnessed during the evaluation, a score of 1 is appropriate. If autism symptoms are better accounted for by another disorder (such as language delay or intellectual disability), a score of 99 is appropriate. If 99 is chosen, please list the disability that better accounts for autism symptoms.

Note that the OARS-4 should only be completed for children who follow a CASE workflow. The CGI should be completed for every child enrolled in SEED.

Scoring OARS-4 and CGI

OARS-4 Domain Impairment Means and Symptom Count: If the subject is nonverbal (impairment in communication, B2), the rater should enter N/A in the right-most box. The impairment means or individual severity ratings are calculated by adding the total number of points gained in each domain and then dividing by the total number of points possible in each domain. The denominators for verbal children are 12 for social interaction, 12 for communication, and 12 for restricted patterns. The denominators for nonverbal children are 12 for social interaction, 9 for communication (excluding B2), and 12 for restricted patterns. The domain symptom counts are calculated by adding the number of symptoms coded 1-3 in each domain. The total number of symptoms possible are 4 for social, 4 for communication, and 4 for restricted patterns; this applies to all children. If the child is nonverbal and B2 is marked N/A the symptom should be treated as present.

OARS-4 Total Impairment Mean and Symptom Count: The total impairment mean (or individual severity rating) is calculated by adding the total number of points gained (overall) and dividing by the total number of points possible (overall). The denominator for verbal children is 36 points and the denominator for nonverbal children is 33 points. The total symptom count is calculated by adding the number of symptoms coded 1-3 in all domains. The total number of symptoms possible is 12 for all children. If the child is nonverbal and B2 is marked N/A the symptom should be treated as present.

OSU Autism CGI scale. These are scored in much the same way as the NIMH precursor of like name. The primary difference between children with severe autism (6) and "classic" autism (7) is that children rated as having "classic" autism have definite associated symptoms (e.g., extreme compulsions, self injury) that are intrusive to others or detrimental to self.

Reference

The OSU Research Unit on Pediatric Psychopharmacology (2005, November): **OSU Autism CGI.** Columbus, OH: Author. © by OSU Research Unit on Pediatric Psychopharmacology, 2005 ; investigators and clinicians may make free copies.

Please place a checkmark for each behavior, based on whether it describes this child never or rarely, sometimes, often, or very often in the past 2 weeks.

	Never or Rarely–Not a Problem (0)	**Sometimes or a little–A Little Problem (1)**	**Often–A Pretty Big Problem (2)**	**Very Often–A Severe Problem (3)**
A: Impairment in social interaction				
1. Impairment in the use of multiple, nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction				
2. Impaired peer relations (compared to developmental level)				
3. Impairment in spontaneous seeking to share enjoyment, interests, or achievements with other people				
4. Impairment in social or emotional reciprocity (returning smiles or greetings, looking at speaker)				
B: Impairment in communication				
1. Does not attempt to speak or communicate; if nonverbal, fails to use gesture or mime to communicate				
2. If adequate speech: impairment in the ability to initiate or sustain a conversation. (Mark N/A in last column if nonverbal.)				
3. Stereotyped and repetitive use of language or sounds or idiosyncratic language				
4. Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level				
C: Restricted repetitive and stereotyped patterns of behavior, interests, and activities				
1. Excessive preoccupation with or stereotyped, restricted patterns of *interest* that are abnormal/odd either in intensity or focus				

2. Inflexible adherence to specific, nonfunctional *routines or rituals*				
3. Stereotyped, repetitive *motor mannerisms* (e.g., hand or finger flapping, or twisting, or complex whole-body movements or "self-stimming")				
4. Preoccupation with parts of objects				

OSU OARS-4 Adapted for SEED Scoring Instructions

Calculate the mean severity rating for each domain and total impairment mean by adding the total number of points gained and then dividing by the total number of points possible. The denominators for verbal children are 12 for social interaction, 12 for communication, 12 for restricted patterns, and 36 for total symptoms. The denominators for nonverbal children are 12 for social interaction, 9 for communication (excluding B2), 12 for restricted patterns, and 33 for total symptoms. Calculate the total number of symptoms present by adding the number of symptoms scored 1-3 in each domain and overall. The number of symptoms possible are 4 for social, 4 for communication, 4 for restricted patterns, and 12 overall (for all children). Remember to count symptom B2 as present if the child is nonverbal.

A: Social Interaction Impairment Mean: __ . __ __ Symptoms Present: __

B: Communication Impairment Mean: __ . __ __ Symptoms Present: __ (If nonverbal, treat symptom B2 present)

C: Restricted Patterns Mean: __ . __ __ Symptoms Present: __

Total Impairment Mean: __ . __ __ (use mean of A, B, & C) Symptoms Present: __ __

2. Curriculum Vitae

Lydia A. Stewart

Phone: (910) 391-2563 | Email: Lstewa25@jhu.edu

EDUCATION

Doctor of Philosophy in Public Health (PhD) **Anticipated, May 2019**
Johns Hopkins Bloomberg School of Public Health, Department of Mental Health
Honors: Children's Mental Health Services Trainee, Wendy Klag Scholar

- Doctoral Dissertation: Association of Socio-demographic Characteristics with Autism Spectrum Disorder: Measurement and Identification Timing in the U.S.

Master of Health Science (MHS) **May 2014**
Johns Hopkins Bloomberg School of Public Health, Department of Mental Health
Honors: Wendy Klag Scholar

- Master's Thesis: *Autism Spectrum Disorder Screening in Low and Middle Income Countries*

Bachelor of Science in Public Health (BSPH) with Highest Honors
May 2013
UNC-Chapel Hill Gillings School of Global Public Health
Department of Health Policy & Management
Honors: AUPHA/McGaw Scholarship

- Honor's Thesis: *Mapping Clergy Response to Parishioner Depression in Rural North Carolina*

WORK EXPERIENCE

Research Associate **Jan. 2016-Present**
Total Child Health/CHADIS

- Child Health and Development Interactive System (CHADIS)—An online screening & clinical decision support platform for pediatric primary care
- *Continuous Quality Improvement (CQI) Research*: Developed and implemented new institutional protocol for conducting CQI research within CHADIS-enrolled pediatric practices. Engaged independently in the following forms of data collection during site-visits to pediatric primary care clients: office workflow mapping, structured interviews and focus groups with clinical and support staff, technological needs assessments, and surveys. Collaborated with account managers and clinical leadership to develop and implement targeted strategies to promote sustained improvement in both CHADIS use and the quality of care delivered.
- *Product Development:* Currently designing a prototype for a data visualization functionality in CHADIS that will allow clinicians and support staff to view individualized statistical process control charts in real-time to better track completion

of recommended screening and clinical decision support tools. Developing a CHADIS-integrated qualitative survey for clinicians enrolled in the TCH-sponsored Maintenance of Certification Program-IV to capture baseline and post-intervention CHADIS-use barriers, targets of intervention, and intervention strategies.

- *Psychometrics Research:* Performed study recruitment, data management, and manuscript writing roles for a multi-site study validating the Quantitative Checklist for Autism in Toddlers (Q-CHAT). Oversaw study recruitment, researcher training, and implementation at the North Carolina study site.

Research Assistant **August 2013-Present**
Wendy Klag Center for Autism and Developmental Disabilities JHSPH

- *Data Collection:* Conduct participant interviews and administer developmental screening tools for national research study examining the role of genetic and environmental exposures in the risk of neuro-developmental disorders: the Study to Explore Early Development (SEED).
- *Psychometrics Analysis:* Currently conducting validation analysis of the Ohio State Autism Rating Scale (OARS) used in the large case-control SEED study with emphasis on the influence of socio-economic variation in instrument performance.

Student Intern **May- Sep. 2014**
Maryland Center for Developmental Disabilities (MCDD)

- Compiled results of prior needs assessment on disability service providers in rural Maryland and drafted report to state funding agencies and community stakeholders.

Summer Intern **May- Aug. 2012**
University of Michigan "Summer Enrichment Program in Health Disparities"
SEP Placement: Blue Care Network

- Developed the template and conceptual model for the BlueCheck Progress Report to increase use of preventative services and impact plan HEDIS score.

SELECTED PUBLICATIONS AND PRESENTATIONS

Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low-and middle-income countries: A systematic review. *Autism*, *21*(5), 527-539.

Wong, Y. S., Yang, C. C., **Stewart, L.,** Chiang, C. H., Wu, C. C., & Iao, L. S. (2018). Use of the Chinese version modified checklist for autism in toddlers in a high-risk sample in Taiwan. *Research in Autism Spectrum Disorders*, *49*, 56-64.

Sturner, R., Howard, B., Bergmann, P., **Stewart, L**., & Afarian, T. E. (2017). Comparison of autism screening in younger and older toddlers. *Journal of Autism and Developmental Disorders*, *47*(10), 3180-3188.

Stewart, L. A., Sturner, R & Howard, B. "Increasing Use of M-CHAT-R/Follow-up in Pediatric Care through Clinician Participation in a Maintenance of Certification Quality Improvement Project". Poster Presented at: International Meeting for Autism Research (IMFAR). May 11, 2017. San Francisco, CA

Stewart, L.A., Lee, L. "Screening for Autism Spectrum Disorders in Low and Middle Income Countries: An Integrative Review". Oral Presentation at: International Meeting for Autism Research (IMFAR). May 14, 2015. Salt Lake City, UT

TEACHING AND MENTORSHIP

Teaching Assistant **2014-Present**
Johns Hopkins Bloomberg School of Public Health

- Facilitated several courses within the Department of Mental Health:
  - *Social, Psychological, and Biological Processes in the Etiology of Mental Disorders (2018)*
  - *Suicide as a Public Health Problem (2015, 2016, 2017)*
  - *Psychiatric Epidemiology (2016)*
  - *Public Mental Health (2016)*
  - *Public Health Perspectives of Autism Spectrum Disorders (2016)*
- Independently developed and delivered interactive lectures to align with course objectives:
  - "Suicide and Suicidal Ideation among Individuals with Autism Spectrum Disorder"
  - "Global Public Mental Health: Challenges and Future Directions"
  - "Screening for Autism Spectrum Disorder in Low and Middle Income Countries"

Volunteer Mentorship **May 2014-Present**

- Mentored several students from Baltimore City Public Schools through design and implementation of student-led original research on topics such as school-based autism services, inclusion of students living with disabilities in extracurricular activities and an overview of the 'behind-the-scenes' work of public health in Baltimore City.
- Provided professional development opportunities based on student-identified goals including: instruction on use of Microsoft Office, developing research posters, one-on-one discussions with prominent scientists at Johns Hopkins, and job interviewing.
- Maintained commitment to improving the diversity and representation of women in science through volunteer work coaching first-generation college students through the undergraduate and graduate school application process and editing application materials.